

BME Matematikai Intézet

Analízis Tanszék

Andai Attila:

Információgeometria a kvantummechanikában

doktori értekezés

Témavezető: Petz Dénes

tudományok doktora, egyetemi tanár

2003



# Tartalomjegyzék

|  |            |
|--|------------|
| <b>Előszó</b>  | <b>iii</b> |
| <b>Bevezetés</b>   | <b>1</b>   |
| <b>1. Statisztikai alapok</b>  | <b>5</b>   |
| 1.1. Fisher-féle információ . . . . .                                      | 5          |
| 1.2. Eloszlások rendezetlensége . . . . .                                  | 14         |
| 1.3. Eloszlások távolsága . . . . .  | 22         |
| 1.4. Majorizáció diszkrét eloszlásoknál . . . . .                          | 27         |
| <b>2. Fisher-metrika a klasszikus esetben</b>                              | <b>31</b>  |
| 2.1. A diszkrét eloszlások Riemann-geometriája . . . . .                   | 32         |
| 2.2. Geodetikusok . . . . .  | 42         |
| 2.3. A térfogat Taylor-sora . . . . .                                      | 46         |
| 2.4. A Fisher-metrika egyértelműsége . . . . .                             | 49         |
| 2.5. Részsokaság skalárgömbülete . . . . .                                 | 52         |
| 2.6. A normális eloszlások geometriája . . . . .                           | 55         |
| 2.7. Az eloszlások geometriai távolsága . . . . .                          | 76         |
| <b>3. Kvantum-információgeometria</b>                                      | <b>81</b>  |
| 3.1. A kvantummechanikai modell . . . . .                                  | 82         |
| 3.2. A kvantummechanikai formalizmus . . . . .                             | 86         |
| 3.3. Állapotok entrópiája és majorizációja . . . . .                       | 88         |
| 3.4. A Fisher-információ általánosítása . . . . .                          | 90         |
| 3.5. Kvantummechanikai Fisher-féle információ . . . . .                    | 95         |
| 3.6. Relatív entrópia . . . . .  | 111        |
| <b>4. Az állapottér geometriája</b>  | <b>121</b> |
| 4.1. Az állapottér skalárgömbülete . . . . .                               | 122        |
| 4.2. A skalárgömbület monotonitásáról . . . . .                            | 140        |
| 4.2.1. Petz sejtésének átfogalmazása . . . . .                             | 140        |
| 4.2.2. Skalárgömbület a Kubo-Mori metrikánál . . . . .                     | 141        |
| 4.2.3. A Petz sejtésében elért eredmények . . . . .                        | 143        |
| 4.2.4. Skalárgömbület monotonitása a valós esetben . . . . .               | 158        |
| 4.3. Az $\mathcal{M}_2^+$ tér skalárgömbülete lokális minimummal . . . . . | 164        |
| 4.4. Az állapottér skalárgömbületének numerikus vizsgálata . . . . .       | 180        |
| 4.5. A térfogat Taylor-sorfejtése . . . . .                                | 189        |
| <b>Jelölésgyűjtemény</b>   | <b>205</b> |

|              |     |
|--------------|-----|
| Tárgymutató  | 209 |
| Hivatkozások | 213 |

## Előszó

Köszönettel tartozom témavezetőmnek, Dr. Petz Dénes Professzornak a munkám során nyújtott segítségéért. A téma kiválasztásában, alapjainak és aktuális eredményeinek a megismerésében, a témán belüli célok kijelölésében fontos szerepe volt. A felmerülő problémák, nehézségek leküzdéséhez sok segítséget kaptam a velem folytatott beszélgetések során. Segítsége nem csak a jelen értekezés elkészítéséhez volt nélkülözhetetlen, hanem az elért új tudományos eredmények szakszerű leírásának a módját is tőle tanultam.

Köszönettel tartozom Dr. Tóth Jánosnak a kézirat gondos átolvasásáért, melynek során számos hibára hívta fel a figyelmemet.

Végül szeretnék köszönetet mondani feleségemnek és szüleimnek, akik támogatásukkal segítettek ezen értekezés megírását.

Budapest, 2003. 11. 13.

*Kijelentem, hogy ezt a doktori értekezést magam készítettem, és abban csak a megadott forrásokat használtam fel. Minden olyan részt, amelyet szó szerint, vagy azonos tartalommal, de átfogalmazva más forrásból átvettem, egyértelműen, a forrás megadásával megjelöltem.*

*A jelen értekezés bírálatai és a védésről készült jegyzőkönyv a későbbiekben a Budapesti Műszaki és Gazdaságtudományi Egyetem Természettudományi Karának Dékáni Hivatalában elérhetők.*



## Bevezetés

Az információgeometria nemkommutatív általánosítását – szórványos kezdeményezésektől eltekintve – az 1990-es években kezdték kidolgozni. Ez a tudományterület, mint ahogy a neve is sugallja, a matematika meglehetősen különböző ágait használja eszközként. Az elsődleges alapot a statisztika és az információelmélet szolgáltatja. Bizonyos statisztikai modellek olyan Riemann-sokaságnak tekinthetők, amelyben a modell differenciálgeometriai jellemzői statisztikai jelentést nyernek. A statisztika és a differenciálgeometria ezen ötvözetét nevezik információgeometriának. Az 1920-as években körvonalazódó kvantummechanika matematikai eszköztárában jelent meg a valószínűségszámítás újfajta, általánosabb megközelítése, melyet ma nemkommutatív valószínűségszámításnak neveznek. A klasszikus statisztikai modell fogalma, mely szorosan kapcsolódik a Kolmogorov-féle valószínűségszámítás fogalmaihoz, általánosítható a nemkommutatív valószínűségszámítás esetére is. Ezt az általánosított statisztikai modellt szintén Riemann-sokasággá lehet tenni. Az így nyert matematikai objektumok képezik a nemkommutatív információgeometria vizsgálatának tárgyát.

A nemkommutatív információgeometria bemutatásához elengedhetetlen a matematika néhány fejezetének a nagyon vázlatos, célorientált, áttekintése. Főként a statisztika, a valószínűségszámítás, a differenciálgeometria és a funkcionálanalízis eszközeit fogjuk alkalmazni speciális esetekben. Ezen eszközök bemutatása során nem célunk a szükséges matematikai fogalmak és tételek legáltalánosabb formájának az áttekintése, megelégszünk azzal a fogalmi kerettel, melyet a nemkommutatív információgeometria kíván. Egy-egy tudományterületen jól ismert tételekhez nem írunk külön-külön referenciát, hanem a fejezetek elején található bevezetőben utalunk az adott tudományterület néhány alapozó jellegű angol, illetve magyar nyelvű irodalmára, ahol a tételek bizonyításokkal és további referenciákkal együtt fellelhetők. A bevezetett alapfogalmakat illetve tételeket gyakran (egymásra épülő) példákon keresztül mutatjuk be. Az előforduló példák végét a  $\blacksquare$  szimbólum jelöli, a bizonyításokét pedig  $\square$ . Az ábrák elkészítése és a numerikus szimulációk elvégzése a Maple matematikai program segítségével történt.

A nemkommutatív információgeometria bizonyos területein elért újabb eredmények bemutatását három, előkészítő jellegű fejezet előzi meg. Az első fejezetben a statisztika és a klasszikus valószínűségszámítás azon alapfogalmait és eredményeit tekintjük át, melyeket a későbbiekben jól tudunk használni a nemkommutatív esetre való általánosításakor. Először a statisztikai modell fogalmát értelmezzük, majd példákon keresztül mutatjuk be a két főbb modelltypust, az exponenciális, illetve a kevert családot. Majd a statisztikai modellhez rendelt Fisher-féle információt definiáljuk, és áttekintjük a főbb tulajdonságait, többek között a paraméterbecslésben központi szerepet játszó Cramer–Rao-tételt. Az entrópia definiálása után röviden

megemlítjük a fizikában sokat használt maximálisentrópia-elv eredetét, és bemutatjuk egy alkalmazását, eljutva így a Gibbs-állapotok fogalmához. Az eloszlások közötti távolság fogalmát, az általánosított divergenciát tisztázzuk ezután, és megmutatjuk, hogy ez szorosan kapcsolódik a Fisher-féle információhoz. Végül az eloszlások rendezetlenségére utaló majorizációs relációt mutatjuk be, és számos ekvivalens feltételt adunk a reláció fennállására.

A második fejezet célja a klasszikus információgeometria alapvető eszközeinek a bemutatása. Néhány példától eltekintve a sokaság mindig valamilyen statisztikai modell lesz a fejezet folyamán. A differenciálgeometriai alapfogalmak áttekintése után közelebbről megvizsgáljuk, hogy milyen kapcsolat van a Riemann-sokaságban lévő gömb térfogatának a sugár szerinti Taylor-sorfejtése és a sokaság görbületét jellemző paraméterek között. Majd megemlítjük Cencov tételét, mely szerint a diszkrét eloszláson alapuló statisztikai modellt statisztikailag releváns módon lényegében csak egyféle képpen lehet Riemann-sokasággá tenni. A részsokaság és a sokaság görbülete között fennálló kapcsolat lehetőséget teremt arra, hogy a statisztikai modellek görbületét egy újabb módszerrel is meghatározhassuk. Ezt közelebbről is megvizsgáljuk a diszkrét- és a normális eloszlás családjának az esetében. Majd egy újabb, geometriai eredetű, távolságfogalmat vezetünk be.

A harmadik fejezet az információgeometria nemkommutatív (kvantummechanikai) általánosításról szól. A kvantummechanika és a matematikai modellje közti kapcsolatot bemutató első rész után a klasszikus valószínűségi eloszlást általánosítjuk, eljutva így a kvantummechanikai állapot fogalmához. Ezen állapotokra is értelmezhető az entrópiafüggvény és a majorizációs reláció, valamint a (kvantummechanikai) maximálisentrópia-elvből származtathatók (a klasszikus esethez hasonlóan) a kvantummechanikai Gibbs-állapotok. A statisztikai modell fogalma is egyszerűen definiálható a nemkommutatív esetben, azonban a Fisher-féle információ általánosítása már messze nem egyértelmű. Bemutatjuk az általánosítás néhány változatát, valamint a Cencov-tétel kvantummechanikai megfelelőjét, a Petz-féle osztályozási tételt, mely szerint az állapottéren értelmezett statisztikailag releváns Riemann-metrikák bizonyos operátor-monoton függvényekkel indexelhetők. Tehát míg a klasszikus esetben egyértelmű a Fisher-féle információ, addig a nemkommutatív esetben sok ilyen Fisher-féle információs mennyiség létezik. Ezek közül az irodalomban leggyakrabban előfordulókat részletesen megvizsgáljuk; és példákon keresztül mutatjuk be, hogy a klasszikus Fisher-féle információhoz kapcsolódó egységes kép, hogyan aprózódik fel a nemkommutatív esetben: például a klasszikus esetben az entrópiafüggvény második deriváltjából is és a megfelelő dimenziójú gömbön értelmezett euklideszi metrikából is egyazon Riemann-metrikát lehetett származtatni a diszkrét statisztikai modelleken, addig a nemkommutatív esetben az első módszer a Kubo–Mori-féle Riemann-metrikát adja, a második pedig a Wigner–Yanase-féle Riemann-metrikát generálja (mely metrikák Petz-osztályozási tétele szerint a klasszikus Fisher-féle információ releváns általánosításai, azon-



---

ban nem azonosak). Megemlítjük továbbá a Cramer–Rao-egyenlőtlenség egyik általánosítását. A relatív entrópiát szintén ki lehet terjeszteni a nemkommutatív esetre, és a klasszikus esethez hasonlóan igazolható, hogy a második deriváltja Fisher-féle információs mennyiséget generál. Ezen relatív entrópiák bizonyos operátorkonvex függvényekkel indexelhetők, és többek között kiderül, hogy kölcsönösen egyértelmű megfeleltetés létezik bizonyos operátorkonvex függvények ekvivalenciaosztályai és a relatív entrópiák által generált Riemann-metrikák között. Továbbá tisztázzuk a nemkommutatív Fisher-féle információ, nemkommutatív relatív entrópia, az operátorkonvex és az operátormonoton függvények egymással való kapcsolatát.

A negyedik fejezetben részletesen megvizsgáljuk a kvantummechanikai állapotok terének néhány differenciálgeometriailag fontos tulajdonságát, abban az esetben, ha Fisher-féle információnak megfelelő Riemann-metrikával látjuk el az állapotteret. Először a sokaság görbületi tenzorát határozzuk meg, valamint a belőle számolható skalárgörbületet. Ezen számítás során segítséget jelent, hogy a második fejezetben már meghatároztuk a többdimenziós normális eloszlások családjának a görbületét. Majd a skalárgörbület kiszámítása után Petz-sejtését, illetve a sejtés bizonyításában eddig elért eredményeket mutatjuk be. A sejtés szerint a kevertebb (vagy kaotikusabb) állapotban nagyobb az állapotter skalárgörbülete, ha az állapotteret a Kubo–Mori-féle Riemann-metrikával látjuk el. Ennek az igazolását nehezíti, hogy a skalárgörbület kifejezése meglehetősen bonyolult formula. Petz-sejtése azonban nem igaz, ha az állapotok terét tetszőleges Fisher-féle Riemann-metrikával láthatjuk el, erre nézünk példát a legegyszerűbb, de még nem triviális kvantumállapotok terén. Majd a bonyolultabb állapottereken numerikus szimulációk segítségével elemezzük a skalárgörbület-függvény viselkedését az irodalomban gyakran előforduló Riemann-metrikák esetében. Végül az állapotterben lévő gömb térfogatának a sugár szerinti Taylor-sorfejtését határozzuk meg bizonyos metrikák esetén.



# 1. Statisztikai alapok

A fejezetben azon statisztikai fogalmakat és alapvető tételeket tekintjük át, melyeket a későbbiekben jól tudunk használni a nemkommutatív valószínűségszámítás területén. A statisztikai definíciókat és tételeket olykor egyszerűsített formában említjük. Ez természetesen megszorítása az általánosabb statisztikai formalizmusnak, azonban a kvantummechanikai (nemkommutatív) esetre való kiterjesztésnél az itt bemutatandó statisztikai fogalmak és tételek bőven elégségesek lesznek. A statisztikából bemutatandó részek megtalálhatóak a bevezető jellegű statisztika könyvek nagy részében, például a magyar nyelvű [16, 17] könyvekben, vagy a további fejezetek olvasásához is segítséget nyújtó [2, 3] könyvekben.

Először a vizsgálódásunk tárgyát a statisztikai modellt, definiáljuk, majd a modell Fisher-féle információját vizsgáljuk meg részletesen. Tételeken keresztül megmutatjuk, hogy a Fisher-féle információs mennyiség miként méri az információt, és milyen szerepe van a statisztikai paraméterbecslésben. Ezt követően a statisztikus fizikában sokat használt maximálisentrópia-elv eredetét, jelentését és példákon keresztül a használatát mutatjuk be. Majd az eloszlások közötti távolságfogalmat tisztázzuk, valamint az ilyen távolságjellegű függvények példáján keresztül igazoljuk, hogy a Fisher-féle információ ezen a területen is központi szerepet játszik. Végül a diszkrét eloszlásokon értelmezett majorizációs relációt mutatjuk be, és ekvivalens feltételeket adunk a reláció fennállására.

## 1.1. Fisher-féle információ

A statisztika és a valószínűségszámítás hatékony eszköz összetett, sokparaméteres rendszerek vizsgálatára. Az elméleti módszer egyik alapfogalma az eloszlásfüggvény. Az adott tudományterület elméleti háttere szolgáltatja a lehetséges eloszlásfüggvények halmazát. A rendszerből vett minták mérésével kiválasztható a vizsgálat szempontjából legmegfelelőbb eloszlásfüggvény az elméletileg adott halmazból. Különböző tudományterületekről származó lehetséges eloszlások halmazának az összefoglaló matematikai általánosítása a statisztikai modell.

**1.1. Definíció.** Az  $\mathcal{S} = (X, \mathcal{B}(X), S, \Xi)$  négyest *statisztikai modellnek* hívjuk, ha

1.  $X$  tetszőleges nem üres halmaz, és  $\mathcal{B}(X)$  az  $X$  halmaz részhalmazainak valamely  $\sigma$ -algebrája,
2.  $S$  elemei valószínűségi mértékek a  $\mathcal{B}(X)$   $\sigma$ -algebrán,
3. létezik egy

$$i : \Xi \rightarrow S \quad \vartheta \mapsto \mu_{\vartheta} \tag{1.1}$$

bijekció.

A továbbiakban csak speciális statisztikai modellekkel foglalkozunk. Az alábbiakat követeljük meg a statisztikai modelltől.

1. Valamilyen pozitív egész  $m$ -re  $\Xi \subseteq \mathbb{R}^m$  teljesül, továbbá  $\Xi$  összefüggő nyílt halmaz. Ekkor azt mondjuk, hogy  $\mathcal{S}$   $m$ -dimenziós statisztikai modell.
2. Ha az  $X$  alaphalmaz véges, akkor azonosítható valamilyen  $n$  természetes szám esetén az  $X_n = \{0, 1, \dots, n\}$  halmazzal. Ekkor  $\mathcal{B}(X)$  az  $X$  hatványhalmaza. Minden  $\vartheta \in \Xi$  esetén  $\mu_\vartheta \in \mathcal{S}$  valószínűségi mérték az  $X_n$  halmazon, melyet azonosítunk a  $p_\vartheta$  sűrűségfüggvényével.
3. Ha az  $X$  alaphalmaz végtelen, akkor azonosítható valamilyen  $n$  természetes szám esetén  $\mathbb{R}^n$  valamely összefüggő, nyílt részhalmazával. Ekkor  $\mathcal{B}(X)$  a tér Borel-féle  $\sigma$ -algebrája. Továbbá minden  $\vartheta \in \Xi$  esetén  $\mu_\vartheta \in \mathcal{S}$  olyan valószínűségi mérték az  $X$  halmazon, melynek létezik  $p_\vartheta$  sűrűségfüggvénye.
4. Az (1.1) leképezés indukál egy topológiát az  $S$  halmazon, ezért  $S$ -et topologikus térnek tekintjük.
5. A 2. és 3. pont miatt az  $S$  halmaz elemeire mint  $X$ -en értelmezett  $(p_\vartheta)_{\vartheta \in \Xi}$  sűrűségfüggvényekre hivatkozunk. Adott  $x \in X$  esetén a  $p(x, \vartheta) = p_\vartheta(x)$  jelölést használjuk.
6. Minden  $p_\vartheta \in \mathcal{S}$  sűrűségfüggvénynek létezik az 1., 2., és 3. momentuma.
7. Minden  $x \in X$  esetén

$$\Xi \rightarrow \mathbb{R} \quad \vartheta \mapsto p(x, \vartheta)$$

végtelen sokszor deriválható leképezés. A  $\Xi \subseteq \mathbb{R}^m$  esetben, a

$$\partial_i p(x, \vartheta) = \frac{\partial p(x, \vartheta)}{\partial \vartheta_i} \quad i = 1, \dots, m \quad (1.2)$$

rövidítést használjuk a  $\vartheta_i$  szerinti parciális deriváltakra.

8. Feltesszük, hogy a  $\vartheta$  szerinti parciális deriválások, és az  $x$  szerinti integrálás sorrendje felcserélhető:

$$\int_X \partial_{i_1} \dots \partial_{i_k} p(x, \vartheta) \, dx = \partial_{i_1} \dots \partial_{i_k} \int_X p(x, \vartheta) \, dx = \partial_{i_1} \dots \partial_{i_k} 1 = 0, \quad (1.3)$$

ahol  $0 < i_1, \dots, i_k \leq m$  a  $\Xi \subseteq \mathbb{R}^m$  esetben.

9. Az  $S$  halmaz elemei szigorúan pozitív függvények, vagyis minden  $\vartheta \in \Xi$  és minden  $x \in X$  esetén  $p(x, \vartheta) > 0$  teljesül.

A fentiek alapján  $X$  egyértelműen meghatározza a  $\mathcal{B}(X)$  halmazrendszert, ezért csak  $(X, S, \Xi)$ -vel jelöljük a statisztikai modellt.

Ezen megszorítások nagy mértékben megkönnyítik a statisztikai modellek differenciálgeometriai analizisét, azonban még mindig elég tág keretet jelentenek a gyakorlatban előforduló főbb modellek számára. Az eloszlások egy részét az alábbi két család valamelyikébe lehet sorolni.

**1.2. Definíció.** Az mondjuk, hogy az  $(X, S, \Xi)$   $n$ -dimenziós statisztikai modell *exponenciális család*, ha léteznek  $C, F_1, \dots, F_n$   $X$ -en értelmezett valós értékű függvények,

$$\tau : \Xi \rightarrow \mathbb{R}^n \quad \xi \mapsto \tau(\xi) \quad (1.4)$$

leképezés és a  $\text{Ran}(\tau)$  halmazon értelmezett  $\psi$  függvény, hogy minden  $p \in S$  felírható

$$p(x, \xi) = \exp \left( C(x) + \sum_{i=1}^n \tau_i(\xi) F_i(x) - \psi(\tau(\xi)) \right) \quad (1.5)$$

alakban. Ekkor a  $\tau$  függvényt *exponenciális-* (vagy *kanonikus-, természetes-*) *paraméterezésnek* nevezzük.

Ha bevezetjük a  $\vartheta_i = \tau_i(\xi)$  és az  $M = \tau(\Xi)$  jelölést, akkor mondhatjuk, hogy az  $S$  statisztikai modell a

$$p(x, \vartheta) = \exp \left( C(x) + \sum_{i=1}^n \vartheta_i F_i(x) - \psi(\vartheta) \right) \quad (1.6)$$

alakú sűrűségfüggvények  $\{p(x, \vartheta) \mid \vartheta \in M\}$  halmaza.

A definíció alapján és a  $\int_X p(x, \vartheta) \, dx = 1$  feltétel miatt

$$\psi(\vartheta) = \log \int_X \exp \left( C(x) + \sum_{i=1}^n \vartheta_i F_i(x) \right) \, dx \quad (1.7)$$

teljesül, valamint  $\{1, F_1, \dots, F_n\}$  lineárisan független függvényhalmaz.

**1.3. Definíció.** Az mondjuk, hogy az  $(X, S, \Xi)$   $n$ -dimenziós statisztikai modell *kevert család*, ha léteznek  $C, F_1, \dots, F_n$   $X$ -en értelmezett valós értékű függvények, hogy minden  $p \in S$  felírható

$$p(x, \vartheta) = C(x) + \sum_{i=1}^n \vartheta_i F_i(x) \quad (1.8)$$

alakban. Ekkor a  $\vartheta \in \Xi$  paraméterezést *kevert paraméterezésnek* nevezzük.

**1.1. Példa.** Diszkrét eloszlás: Az alaphalmaz  $X = \{0, 1, \dots, n\}$ , a paraméterter

$$\Xi = \left\{ (\eta_1, \dots, \eta_n) \in \mathbb{R}^n \mid \forall i \in \{1, \dots, n\} : 0 < \eta_i < 1, \sum_{i=1}^n \eta_i < 1 \right\}. \quad (1.9)$$

Legyen  $S$  az  $X$  halmazon értelmezett minden  $x \in X$  elemen szigorúan pozitív értéket felvevő sűrűségfüggvények halmaza. Ekkor  $(X, S, \Xi)$  statisztikai modell. Bevezetve a

$$C^{(\eta)}(x) = \begin{cases} 1 & \text{ha } x = 0 \\ 0 & \text{ha } x \neq 0, \end{cases} \quad \forall i \in \{1, \dots, n\} : F_i^{(\eta)}(x) = \begin{cases} -1 & \text{ha } x = 0 \\ 1 & \text{ha } x = i \\ 0 & \text{egyébként} \end{cases} \quad (1.10)$$

a függvényeket, látható, hogy a diszkrét eloszlások a kevert családba tartoznak, mivel az  $S$  halmaz a

$$p(x, \eta) = C^{(\eta)}(x) + \sum_{i=1}^n \eta_i F_i^{(\eta)}(x) \quad (1.11)$$

sűrűségfüggvények

$$\{p(x, \eta) \mid \eta \in \Xi\}$$

halmazával egyezik meg.

Vezessük be a

$$\vartheta_i(\eta) = \log \frac{\eta_i}{1 - \sum_{k=1}^n \eta_k} \quad (1.12)$$

mennyiségeket. Ekkor minden  $i \in X$  és  $\eta \in \Xi$  esetén

$$p(i, \eta) = \frac{e^{\vartheta_i}}{1 + \sum_{k=1}^n e^{\vartheta_k}} \quad (1.13)$$

teljesül. Bevezetve a

$$C^{(\vartheta)}(x) = 0, \quad \psi(\vartheta) = \log \left( 1 + \sum_{i=1}^n e^{\vartheta_i} \right), \quad F_i^{(\vartheta)}(x) = \begin{cases} 1 & \text{ha } x = i \\ 0 & \text{ha } x \neq i \end{cases} \quad (1.14)$$

függvényeket minden  $p(x, \eta) \in S$  felírható

$$p(x, \eta) = \exp \left( C^{(\vartheta)}(x) + \sum_{i=1}^n \vartheta_i(\eta) F_i^{(\vartheta)}(x) - \psi(\vartheta(\eta)) \right) \quad (1.15)$$

alakban. Ezek szerint az  $(X, S, \Xi)$  statisztikai modell az exponenciális családba tartozik a  $\vartheta$  kanonikus paraméterekkel.  $\blacksquare$

**1.2. Példa.** Normális eloszlás: Az alaphalmaz  $X = \mathbb{R}$ , a paraméterter  $\Xi = \mathbb{R} \times \mathbb{R}^+$ . Az  $S$  halmaz elemei a

$$p(x, \mu, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x - \mu)^2}{2\sigma^2}\right) \quad (1.16)$$

sűrűségfüggvények. Legyen

$$\tau : \Xi \rightarrow \mathbb{R} \times \mathbb{R}^- \quad (\mu, \sigma) \mapsto (\vartheta_1, \vartheta_2) = \left(\frac{\mu}{\sigma^2}, -\frac{1}{2\sigma^2}\right) \quad (1.17)$$

leképezés, valamint  $C(x) = 0$ ,  $F_1(x) = x$  és  $F_2(x) = x^2$ . A

$$\psi(\vartheta_1, \vartheta_2) = \frac{-\vartheta_1^2}{4\vartheta_2} + \frac{1}{2} \log\left(-\frac{\pi}{\vartheta_2}\right) \quad (1.18)$$

függvénnyel a normális eloszlás az exponenciális családba tartozik a  $(\vartheta_1, \vartheta_2)$  kanonikus paraméterekkel.

Hasonló paraméterezéssel igazolható, hogy a többdimenziós normális eloszlás is az exponenciális családba tartozik.  $\square$

Egy  $(X, S, \Xi)$   $n$ -dimenziós statisztikai modell esetén a Fisher-féle információ minden  $\vartheta \in \Xi$  paraméternél egy  $n \times n$ -es mátrixot ad értékül.

**1.4. Definíció.** Legyen  $(X, S, \Xi)$   $n$ -dimenziós statisztikai modell. Adott  $\vartheta \in \Xi$  pont esetén a *Fisher-féle információs mátrix* az az  $n \times n$ -es mátrix, melynek  $(i, k)$ -adik eleme

$$g^{(F)}(\vartheta)_{ik} = \int_X \frac{1}{p(x, \vartheta)} (\partial_i p(x, \vartheta)) (\partial_k p(x, \vartheta)) \, dx. \quad (1.19)$$

A későbbiekben  $g^{(F)}(\vartheta)$  jelöli a Fisher-féle információs mátrixot. A definíció alapján közel sem nyilvánvaló, hogy miért és miként mér információt a  $g^{(F)}(\vartheta)$  mátrix. Ennek egyik oka a formula nehezen értelmezhető szerkezete, a másik pedig, hogy az információ adott tudományterülethez tartozó mértékegységét, mérésének a módját nem egyszerű meghatározni. A  $g^{(F)}(\vartheta)$  mátrix tulajdonságai és a rá vonatkozó tételek viszont alapot adnak arra, hogy egyfajta információs mennyiségként értelmezzük. A Fisher-féle információs mátrixra az alábbi reprezentációkat fogjuk még használni:

$$g^{(F)}(\vartheta)_{ik} = \int_X p(x, \vartheta) (\partial_i \log p(x, \vartheta)) (\partial_k \log p(x, \vartheta)) \, dx \quad (1.20)$$

$$g^{(F)}(\vartheta)_{ik} = 4 \int_X (\partial_i \sqrt{p(x, \vartheta)}) (\partial_k \sqrt{p(x, \vartheta)}) \, dx. \quad (1.21)$$

**1.1. Tétel.** Legyen  $(X, S, \Xi)$   $n$ -dimenziós statisztikai modell. Azokban a  $\vartheta \in \Xi$  pontokban, ahol a  $(\partial_i p(\cdot, \vartheta))_{i=1, \dots, n}$  függvények lineárisan függetlenek, a Fisher-féle  $g^{(F)}(\vartheta)$  információs mátrix pozitív definit.

*Bizonyítás.* Legyen  $(c_1, \dots, c_n) \in \mathbb{R}^n$  tetszőleges vektor. Ekkor

$$\langle (c_1, \dots, c_n), g^{(F)}(\vartheta)(c_1, \dots, c_n) \rangle = \int_X p(x, \vartheta) \left( \sum_{i=1}^n c_i \partial_i (\log p(x, \vartheta)) \right)^2 dx \geq 0. \quad (1.22)$$

A fenti egyenlőtlenség alapján  $g^{(F)}(\vartheta)$  pozitív szemidefinit. A tételben szereplő feltétel pedig garantálja, hogy a fenti egyenlőtlenségben az integrál után szereplő négyzet legalább egy  $x$  helyen pozitív legyen. A folytonossági feltételek miatt az integrál határozottan pozitív.  $\square$

Egy mérésekben rejlő információ mértékére természetes feltétel annak a megkövetelése, hogy ugyanazt a jelenséget mérve jobb, illetve rosszabb felbontású műszerrel, a rosszabb műszerrel való mérés esetén ne lehessen több információhoz jutni, mint a jobb műszerrel. Például, ha egy dobókockát vizsgálunk, akkor több információhoz jutunk, ha látjuk a dobott számot, mint ha csak azt tudjuk, hogy párosat vagy páratlant dobtunk. Ennek a természetes feltételnek a Fisher-féle információ eleget tesz.

Legyen  $(X, S, \Xi)$  statisztikai modell és

$$f : X \rightarrow Y \quad x \mapsto f(x) \quad (1.23)$$

szürjektív leképezés. Az  $f$  mérhető függvénnyel minden  $p \in S$  függvényt áttranszformálhatunk egy

$$\tilde{p} : Y \rightarrow \mathbb{R} \quad y \mapsto \tilde{p}(y, \vartheta) \quad (1.24)$$

függvénybe a

$$\forall A \in \mathcal{B}(X) : \int_A p(x, \vartheta) dx = \int_{f(A)} \tilde{p}(y, \vartheta) dy \quad (1.25)$$

összefüggés segítségével. Véges  $X$  halmaz esetén

$$\tilde{p}(y, \vartheta) = \sum_{x \in f^{-1}(y)} p(x, \vartheta) \quad (1.26)$$

a transzformációs képlet.

Jelölje  $\tilde{f}$  a  $p \rightarrow \tilde{p}$  leképezést. A  $Q = \tilde{f}(S)$  jelöléssel a  $(Y, Q, \Xi)$  hármas statisztikai modell, ez az  $(X, S, \Xi)$  modell  $f$  általi képe.



**1.5. Definíció.** Legyen  $(X, S, \Xi)$  statisztikai modell és  $f : X \rightarrow Y$  szürjektív leképezés. Tekintsük az

$$r(\cdot, \cdot) : X \times \Xi \rightarrow \mathbb{R} \quad (x, \vartheta) \mapsto r(x, \vartheta) = \frac{p(x, \vartheta)}{\tilde{p}(f(x), \vartheta)} \quad (1.27)$$

függvényt. Azt mondjuk, hogy  $f$  *elégséges statisztikája* az  $S$  modellnek, ha minden  $x \in X$  esetén az

$$r(x, \cdot) : \Xi \rightarrow \mathbb{R} \quad \vartheta \mapsto r(x, \vartheta)$$

függvény állandó.

Ha a definícióban szereplő  $f$  függvény bijekció, akkor  $f$  elégséges statisztikája az  $S$  modellnek.

**1.2. Tétel.** Legyen  $(X, S, \Xi)$  statisztikai modell,  $f : X \rightarrow Y$  szürjektív leképezés és  $(Y, Q, \Xi)$  az  $f$  által indukált modell. A Fisher-féle információs mátrix  $\vartheta \in \Xi$  paraméter esetén legyen az  $S$  modellen  $g_S^{(F)}(\vartheta)$  és a  $Q$  modellen  $g_Q^{(F)}(\vartheta)$ . Ekkor minden  $\vartheta \in \Xi$  esetén

$$g_Q^{(F)}(\vartheta) \leq g_S^{(F)}(\vartheta) \quad (1.28)$$

teljesül. A  $\Delta g(\vartheta) = g_S^{(F)}(\vartheta) - g_Q^{(F)}(\vartheta)$  információs veszteség

$$\Delta g_{ik}(\vartheta) = \int_X p(x, \vartheta) \frac{\partial \log r(x, \vartheta)}{\partial \vartheta_i} \frac{\partial \log r(x, \vartheta)}{\partial \vartheta_k} dx. \quad (1.29)$$

Az egyenlőség a (1.28) pontban pontosan akkor teljesül, ha  $f$  elégséges statisztikája az  $S$  modellnek.

A tételben szereplő  $g_Q^{(F)}(\vartheta) \leq g_S^{(F)}(\vartheta)$  egyenlőtlenség jelentése, hogy a  $g_S^{(F)}(\vartheta) - g_Q^{(F)}(\vartheta)$  mátrix pozitív szemidefinit. A tételt gyakran a *Fisher-féle információ monotonitásaként* említik.

A fenti tétel kvantummechanikai esetre való általánosításánál nehézséget jelent, hogy a kvantumos esetben nincs jól értelmezhető  $X$  alaphalmaz. Ezért szükség lesz a fenti tétel átmenetvalószínűségekre vonatkozó megfogalmazására.

**1.6. Definíció.** Legyen  $X \subseteq \mathbb{R}^n$  és  $Y \subseteq \mathbb{R}^m$  nyílt halmaz. Azt mondjuk, hogy a

$$\kappa : Y \times X \rightarrow \mathbb{R} \quad (x, y) \mapsto \kappa(y|x) \quad (1.30)$$

függvény az  $X$  halmazon értelmezett *Markov-féle magfüggvény* vagy *átmenetvalószínűség*, ha minden  $x \in X$  és  $y \in Y$  esetén  $\kappa(y|x) \geq 0$ , továbbá minden  $x \in X$  esetén

$$\int_Y \kappa(y|x) dy = 1 \quad (1.31)$$

teljesül.

Az  $X$  és  $Y$  véges halmazok esetén  $\kappa$  akkor átmenetvalószínűség, ha minden  $x \in X$  esetén

$$\sum_{y \in Y} \kappa(y|x) dy = 1 \quad (1.32)$$

teljesül.

**1.3. Tétel.** Legyen  $(X, S, \Xi)$  statisztikai modell,  $Y$  tetszőleges halmaz és

$$\kappa : Y \times X \rightarrow \mathbb{R} \quad (y, x) \mapsto \kappa(y|x) \quad (1.33)$$

átmenetvalószínűség. A  $\tilde{p}(y, \vartheta) = \int_X \kappa(y|x)p(x, \vartheta) dx$  alakú sűrűségfüggvényekből álló statisztikai modell legyen  $(Y, Q, \Xi)$ . Ekkor minden  $\vartheta \in \Xi$  esetén

$$g_Q^{(F)}(\vartheta) \leq g_S^{(F)}(\vartheta) . \quad (1.34)$$

A  $\Delta g(\vartheta) = g_S^{(F)}(\vartheta) - g_Q^{(F)}(\vartheta)$  információs veszteség pedig

$$\Delta g_{ik}(\vartheta) = \int_X p(x, \vartheta) \frac{\partial \log r(x, \vartheta)}{\partial \vartheta_i} \frac{\partial \log r(x, \vartheta)}{\partial \vartheta_k} dx . \quad (1.35)$$

Az (1.2.) tétel egyszerű következménye a Fisher-féle információ additivitása.

Ennek intuitív jelentése az, hogy ha két egymással nem kölcsönható részből álló rendszer tagjain végzünk független méréseket, akkor az összetett rendszerről nyert információ megegyezik a részekből nyert információk összegével.

**1.7. Definíció.** Legyen  $(X, S, \Xi)$  és  $(Y, Q, \Xi)$  statisztikai modell. Jelölje  $S \times_* Q$  a

$$p : X \times Y \times \Xi \rightarrow \mathbb{R} \quad (x, y, \vartheta) \mapsto p(x, y, \vartheta) = p_1(x, \vartheta)p_2(y, \vartheta) \quad (1.36)$$

alakú paraméteres sűrűségfüggvények halmazát, ahol  $p_1(x, \vartheta) \in S$  és  $p_2(y, \vartheta) \in Q$ . Ekkor az  $(X \times Y, S \times_* Q, \Xi)$  hármast az eredeti *statisztikus modellek független szorzatának* nevezzük. Az  $(X, S, \Xi)$  statisztikai modell önmagával vett  $n$ -szeres független szorzatát  $(X^n, S^{(n)}, \Xi)$ -vel jelöljük. Adott  $p(x, \vartheta) \in S$  esetén a

$$p^{(n)} : X^n \times \Xi \rightarrow \mathbb{R} \quad (x_1, \dots, x_n, \vartheta) \mapsto p^{(n)}(x_1, \dots, x_n, \vartheta) = p(x_1, \vartheta) \cdots p(x_n, \vartheta) \quad (1.37)$$

jelölést használjuk.

**1.4. Tétel.** Legyen  $(X, S, \Xi)$  és  $(Y, Q, \Xi)$  statisztikai modell a  $g_S^{(F)}(\vartheta)$  és a  $g_Q^{(F)}(\vartheta)$  Fisher-féle információs mátrixszal. Jelölje  $g_{SQ}^{(F)}(\vartheta)$  a  $(X \times Y, S \times_* Q, \Xi)$  statisztikai modell Fisher-féle információs mátrixát. Ekkor

$$g_{SQ}^{(F)}(\vartheta) = g_S^{(F)}(\vartheta) + g_Q^{(F)}(\vartheta) \quad (1.38)$$

teljesül minden  $\vartheta \in \Xi$  elemre.

A Fisher-féle információ központi szerepet játszik a statisztikai paraméterbecslésben. Tegyük fel, hogy valamilyen rendszert a  $(X, S, \Xi)$  statisztikai modell jellemez. A rendszeren végzett mérések alapján szeretnénk megbecsülni a rendszer állapotát jellemző  $p \in S$  sűrűségfüggvényben szereplő  $\vartheta$  paramétert. A rendszeren végzett mérést leíró

$$\mathfrak{X} : X \rightarrow \mathbb{R}^m \quad (1.39)$$

függvényt *mintának* nevezzük. Az  $m = 1$  esetben számértékű a mérés végeredménye, egyébként vektorértékű. A rendszeren végzett  $k$  számú mérés alapján a rendszer állapotát meghatározó  $\vartheta$  paraméter becslését egy

$$\tilde{\vartheta} : (\mathbb{R}^m)^k \rightarrow \Xi \quad (x_1, \dots, x_k) \mapsto \tilde{\vartheta}(x_1, \dots, x_k) \quad (1.40)$$

függvény, a *statisztika* adja meg. A  $\tilde{\vartheta}$  függvényt gyakran *becslésnek* vagy *paraméterbecslésnek* is nevezzük. Ha feltesszük, hogy a mérések egymástól függetlenek, akkor az  $x_i$  mennyiségek független azonos eloszlású valószínűségi változók lesznek. Jelölje  $E_{\vartheta}(\tilde{\vartheta})$  a  $\tilde{\vartheta}$  várható értékét a  $p^{(k)}(x, \vartheta)$  eloszlásra nézve, vagyis

$$E_{\vartheta}(\tilde{\vartheta}) = \int_{X^k} p^{(k)}(x, \vartheta) \tilde{\vartheta}(x) \, dx. \quad (1.41)$$

A  $\tilde{\vartheta}$  statisztikáról azt mondjuk, hogy *torzítatlan*, ha minden  $\vartheta \in \Xi$  esetén

$$E_{\vartheta}(\tilde{\vartheta}) = \vartheta \quad (1.42)$$

teljesül. A

$$\begin{aligned} V_{\vartheta}(\tilde{\vartheta})_{ij} &= E_{\vartheta}((\tilde{\vartheta} - E_{\vartheta}(\tilde{\vartheta}))_i (\tilde{\vartheta} - E_{\vartheta}(\tilde{\vartheta}))_j) = \\ &= \int_{X^k} p^{(k)}(x, \vartheta) (\tilde{\vartheta}(x) - E_{\vartheta}(\tilde{\vartheta}))_i (\tilde{\vartheta}(x) - E_{\vartheta}(\tilde{\vartheta}))_j \, dx \end{aligned} \quad (1.43)$$

$k \times k$  mátrixot a *statisztika varianciájának* nevezzük. A torzítatlan becslés varianciájára létezik egy alsó küszöb, mely a Fisher-féle információval fejezhető ki, ezt fogalmazza meg az alábbi tétel.

**1.5. Tétel.** *Cramer–Rao: Legyen  $(X, S, \Xi)$  statisztikai modell,  $k$  tetszőleges természetes szám, jelölje  $g^{(F)}$  a  $(X^k, S^{(k)}, \Xi)$  statisztikai modell Fisher-féle információját, legyen  $\tilde{\vartheta}$  torzítatlan becslése a  $\vartheta$  paraméternek, és jelölje  $V_{(\vartheta)}(\tilde{\vartheta})$  a becslés varianciáját. Ekkor minden  $\vartheta \in \Xi$  esetén*

$$V_{\vartheta}(\tilde{\vartheta}) \geq (g^{(F)}(\vartheta))^{-1} \quad (1.44)$$

teljesül.

## 1.2. Eloszlások rendezetlensége

Egy  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  sűrűségfüggvénynek is értelmezhető a Fisher-féle információja természetes módon. Definiáljuk a

$$\tilde{f} : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R} \quad (x, y) \mapsto \tilde{f}(x, y) = f(x + y) \quad (1.45)$$

paraméteres sűrűségfüggvényt. Ennek a Fisher-féle információja

$$g_{ik}(y) = \int_{\mathbb{R}^n} \frac{1}{\tilde{f}(x, y)} \frac{\partial \tilde{f}(x, y)}{\partial y_i} \frac{\partial \tilde{f}(x, y)}{\partial y_k} dx. \quad (1.46)$$

Ez a mennyiség független  $y$ -től, tehát a Fisher-féle információ ekkor a

$$g_{ik} = \int_{\mathbb{R}^n} \frac{1}{p(x)} \frac{\partial p(x)}{\partial x_i} \frac{\partial p(x)}{\partial x_k} dx \quad (1.47)$$

mátrix.

Az információnak, illetve a bizonytalanságnak másik fontos jellemzője a Boltzmann-féle entrópia.

**1.8. Definíció.** Adott  $f : X \rightarrow \mathbb{R}$  sűrűségfüggvény *entrópiája*

$$S(f) = - \int_X f(x) \log f(x) dx, \quad (1.48)$$

ha az integrál létezik és véges. Az  $f(x) = 0$  esetben a  $0 \log 0 = 0$  konvenció használandó.

**1.3. Példa.** Normális eloszlás: Az

$$f(x) = \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad (1.49)$$

sűrűségfüggvényű normális eloszlás Fisher-féle információja

$$g^{(F)} = \int_{\mathbb{R}} \frac{1}{f(x)} \left( \frac{df(x)}{dx} \right)^2 dx = \frac{1}{\sigma^2}, \quad (1.50)$$

entrópiája pedig

$$S(f) = \log \sigma + \log \sqrt{2\pi e}. \quad (1.51)$$

Mivel az entrópia és az abból származtatott mennyiségek központi szerepet fognak játszani a későbbiekben, érdemes röviden megemlíteni, hogy milyen fizikai kontextusban vezették be az entrópia fogalmát.

A fizikában jóval Shannon előtt Boltzmann vezetett be egy  $H$  funkcionált, melyet entrópiának nevezett, és amelynek képlete formálisan megegyezik a Shannon-entrópiával. (Miután az információ mérésében Shannon hasznosnak találta a  $-\sum_i p_i \log p_i$  mennyiséget, Neumann tanácsára adta neki szintén az entrópia nevet. Precízen külön kellene Boltzmann- illetve Shannon-entrópiáról beszélni.) 1872-ben Boltzmann levezette a róla elnevezett transzportegyenletet, mely a gázban lévő molekulák eloszlására vonatkozik. (Nemrelativisztikus kinetikával mozgó és azonos molekulákból álló gázra vonatkozik, ahol elhanyagolható a molekulák között ható erő.) Az  $\frac{1}{N} f(\underline{r}, \underline{v}, t)$  függvény az  $N$  molekulából álló gázban a molekulák számának  $\underline{r}$  helytől,  $\underline{v}$  sebességtől és  $t$  időtől függő eloszlásának a sűrűségfüggvénye. Adott  $\underline{u}, \underline{v} \in \mathbb{R}^3$  vektorok esetén jelölje  $\alpha(\underline{u}, \underline{v})$  a vektorok által bezárt szöveget,  $\sigma(\alpha)$  a molekulák ütközésénél fellépő differenciális hatáskeresztmetszetet,  $\underline{F}$  a molekulákra ható külső erőt,  $m$  egy molekula tömegét és  $N$  a molekulák számát. Ekkor

$$\begin{aligned} \frac{\partial f(\underline{r}, \underline{v}_1, t)}{\partial t} + \underline{v}_1 \text{grad} f(\underline{r}, \underline{v}_1, t) + \frac{\underline{F}}{m} \text{grad}_{\underline{v}_1} f(\underline{r}, \underline{v}_1, t) \\ = N \int \dots \int \sigma(\alpha(\underline{v}_2 - \underline{v}_1, \underline{v}'_2 - \underline{v}'_1)) \cdot |\underline{v}_1 - \underline{v}_2| \cdot \\ \cdot \left( f(\underline{r}, \underline{v}'_1, t) f(\underline{r}, \underline{v}'_2, t) - f(\underline{r}, \underline{v}_1, t) f(\underline{r}, \underline{v}_2, t) \right) d\alpha d\underline{v}_2 \end{aligned} \quad (1.52)$$

a Boltzmann-féle transzportegyenlet, ahol a  $\underline{v}_1, \underline{v}_2$  sebességgel teljesen rugalmasan ütköző molekulák sebessége  $\underline{v}'_1$  illetve  $\underline{v}'_2$  az ütközés után. Tegyük fel, hogy a gázra külső erő nem hat ( $\underline{F} = 0$ ). Ebben az esetben az  $f$  eloszlásfüggvény független a helytől és az

$$\begin{aligned} \frac{\partial f(\underline{v}_1, t)}{\partial t} = N \int \dots \int \sigma(\alpha(\underline{v}_2 - \underline{v}_1, \underline{v}'_2 - \underline{v}'_1)) \cdot |\underline{v}_1 - \underline{v}_2| \cdot \\ \cdot \left( f(\underline{v}'_1, t) f(\underline{v}'_2, t) - f(\underline{v}_1, t) f(\underline{v}_2, t) \right) d\alpha d\underline{v}_2 \end{aligned} \quad (1.53)$$

egyenleteknek tesz eleget. A gáz állapota az egyensúlyi esetben nem változik. Az egyensúlyt jellemző  $f_0$  eloszlás az

$$0 = \int \dots \int \sigma(\alpha(\underline{v}_2 - \underline{v}_1, \underline{v}'_2 - \underline{v}'_1)) \cdot |\underline{v}_1 - \underline{v}_2| \cdot (f_0(\underline{v}'_1) f_0(\underline{v}'_2) - f_0(\underline{v}_1) f_0(\underline{v}_2)) d\alpha d\underline{v}_2 \quad (1.54)$$

egyenletből határozható meg. Vezessük be a

$$H_f(t) = \int f(\underline{v}, t) \log f(\underline{v}, t) d\underline{v} \quad (1.55)$$

funkcionált. Ekkor az alábbiakat állíthatjuk.

1. A

$$\frac{dH_f(t)}{dt} = 0 \iff f = f_0 \quad (1.56)$$

ekvivalencia teljesül.

2. Minden  $f$  eloszlás esetén

$$\frac{dH_f(t)}{dt} \leq 0. \quad (1.57)$$

3. Fennáll a

$$\lim_{t \rightarrow \infty} f(\underline{v}, t) = f_0(\underline{v}) \quad (1.58)$$

reláció.

A fenti állításokat összefoglalóan Boltzmann-féle  $H$ -tételként említik. Az 1. és a 2. pont adja az alapját a termodinamika második főtételenek, melynek egyik megfogalmazása szerint a magára hagyott rendszer entrópiája nem csökken. A 3. pont megegyezik azon fizikai tapasztalattal, mely szerint a kezdetben rendezetlen gázban idővel kialakul az egyensúlyi (maximális entrópiájú) helyzet. Ezek az észrevételek adják az alapját a *maximális entrópia elvének*, mely szerint ha egy rendszer bizonyos mérhető paraméterei többféle eloszlásfüggvényből származhatnak, akkor a rendszer (már kialakult egyensúlyi) állapotát az az eloszlásfüggvény írja le, amelynek az entrópiája maximális.

**1.4. Példa.** Gibbs-eloszlás: Tegyük fel, hogy egy rendszerben  $n$ -féle energiaállapot van  $(E_i)_{1 \leq i \leq n}$  pozitív energiaszintekkel, és léteznek különböző nagyságú energiaszintek. Tegyük fel továbbá, hogy a rendszerben lévő részecskék átlagos energiája  $E$ , és ez az érték a legnagyobb, illetve a legkisebb energiaszint között van, vagyis

$$\min_{1 \leq i \leq n} E_i < E < \max_{1 \leq i \leq n} E_i \quad (1.59)$$

teljesül. Kérdés a részecske energiaszintenkénti eloszlása, vagyis a  $(p_i)_{1 \leq i \leq n}$  értékek, ahol  $p_i$  jelöli a részecske  $E_i$  energiaszinten való tartózkodásának a valószínűségét.

A maximális entrópia elve szerint az

$$S : \mathbb{R}^n \rightarrow \mathbb{R} \quad (p_1, \dots, p_n) \mapsto - \sum_{i=1}^n p_i \log p_i \quad (1.60)$$

függvény maximumát keressük az  $\sum_{i=1}^n p_i = 1$  és a  $\sum_{i=1}^n E_i p_i = E$  feltételek mellett.

A feltételes szélsőérték tétele kimondja, hogy ha  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  folytonosan differenciálható függvény,  $g_0, \dots, g_l : \mathbb{R}^n \rightarrow \mathbb{R}$  folytonosan differenciálható függvények, és az  $f$  függvénynek lokális szélsőértéke van a

$$G = \bigcap_{i=0}^l g_i^{-1}(0) \quad (1.61)$$

halmaz  $a \in G$  pontjában, továbbá, ha a  $(g'_i(a))_{i=0, \dots, l}$  vektorok lineárisan függetlenek, akkor léteznek egyértelműen olyan  $(\lambda_i)_{i=0, \dots, l}$  számok, hogy

$$f'(a) = \sum_{i=0}^l \lambda_i g'_i(a) \quad (1.62)$$

teljesül.

Az  $S$  entrópiára, valamint a

$$g_0 : \mathbb{R}^n \rightarrow \mathbb{R} \quad (p_1, \dots, p_n) \mapsto \sum_{i=1}^n p_i - 1 \quad (1.63)$$

$$g_1 : \mathbb{R}^n \rightarrow \mathbb{R} \quad (p_1, \dots, p_n) \mapsto \sum_{i=1}^n E_i p_i - E$$

függvényekre alkalmazva a feltételes szélsőérték tételét, azt mondhatjuk, hogy a  $(p_1, \dots, p_n)$  pontban, akkor lehet feltételes szélsőértéke  $S$ -nek, ha léteznek olyan  $\lambda_0, \beta$  számok, hogy

$$(-\log p_1 - 1, \dots, -\log p_n - 1) = \lambda_0(1, \dots, 1) + \beta(E_1, \dots, E_n) \quad (1.64)$$

teljesül. A  $g_0(p_1, \dots, p_n) = 0$  feltétel alapján a

$$\lambda_0 = \log \left( \sum_{k=1}^n e^{-\beta E_k} \right) - 1 \quad (1.65)$$

egyenletet kapjuk. Vagyis minden  $i = 1, \dots, n$  esetén

$$p_i = \frac{e^{-\beta E_i}}{\sum_{k=1}^n e^{-\beta E_k}} \quad (1.66)$$

alakba írható. A  $g_1(p_1, \dots, p_n) = 0$  feltételből határozható meg  $\beta$  értéke. Ehhez pozitív  $k$  értékekre definiáljuk az

$$M(E^k) : \mathbb{R} \rightarrow \mathbb{R} \quad \beta \mapsto \frac{\sum_{i=1}^n E_i^k e^{-\beta E_i}}{\sum_{i=1}^n e^{-\beta E_i}} \quad (1.67)$$

függvényt, mely az energia  $(p_0, \dots, p_n)$  eloszláshoz tartozó  $k$ -adik momentuma. Ekkor  $M(E)$  az energia várható értéke. A

$$\lim_{\beta \rightarrow -\infty} M(E)(\beta) = \max_{1 \leq i \leq n} E_i \quad \lim_{\beta \rightarrow \infty} M(E)(\beta) = \min_{1 \leq i \leq n} E_i \quad (1.68)$$

határértékek és a  $M(E)$  függvény monoton fogyását igazoló

$$\frac{dM(E)(\beta)}{d\beta} = M(E)(\beta)^2 - M(E^2)(\beta) \leq 0 \quad (1.69)$$

egyenlőtlenség alapján pontosan egy olyan  $\beta$  paraméter lesz, mely kielégíti a  $g_1$  feltételt. Hátra van még annak igazolása, hogy az (1.66) képlettel értelmezett valószínűségek esetén az entrópiának maximuma van. Legyen  $(q_1, \dots, q_n)$  olyan eloszlás, mely teljesíti a  $g_0$  és  $g_1$  feltételt. Ekkor bevezetve az  $r_i = \frac{q_i}{p_i}$  mennyiségeket az

$$S(p_1, \dots, p_n) - S(q_1, \dots, q_n) = \sum_{i=1}^n p_i (r_i \log r_i - r_i + 1) \quad (1.70)$$

összefüggés, valamint a pozitív számokon értelmezett  $x \mapsto x \log x - x + 1$  függvény pozitivitása miatt kapjuk, hogy a (1.66) képlettel megadott  $(p_i)_{1 \leq i \leq n}$  eloszlás esetén lesz az entrópiának maximuma.

Fizikai alkalmazásokban  $\beta = \frac{1}{kT}$ , ahol  $k$  jelöli a Boltzmann állandót és  $T$  az abszolút hőmérsékletet. Az (1.67)-ben definiált  $M(E)$  függvény monotonitása és  $\beta = 0$  helyen felvett értéke alapján a következőt érdemes megfigyelni. A  $T$  hőmérséklet pontosan akkor pozitív, ha

$$\frac{1}{n} \sum_{i=1}^n E_i < E, \quad (1.71)$$

pontosan akkor negatív, ha

$$\frac{1}{n} \sum_{i=1}^n E_i > E, \quad (1.72)$$

és pontosan akkor nem értelmezhető (végtelen), ha

$$\frac{1}{n} \sum_{i=1}^n E_i = E \quad (1.73)$$

teljesül. Ezen megfigyelés alapján egy rendszernek lehet negatív hőmérséklete, melyet nem az (elérhetetlen) abszolút zérus ponton való túlhűtéssel, hanem a végtelen nagy hőmérsékleten keresztül lehet elérni. (Ezt igazolja, hogy  $T_1 < 0$ ,  $T_2 > 0$  hőmérsékletű rendszerek kölcsönhatása esetén a kialakult közös  $T$  hőmérsékletre  $T < T_1$  vagy  $T > T_2$



teljesül.) Ilyen negatív hőmérsékletű rendszereket fizikailag elő lehet állítani, bővebben lásd a [62] könyv 73. fejezetét.

Érdeemes megemlíteni, hogy a  $(p_1, \dots, p_n)$  eloszlás entrópiájára

$$\frac{\partial S(\beta)}{\partial \beta} = -\beta(M(E^2)(\beta) - M(E)(\beta)^2) \quad (1.74)$$

teljesül, vagyis pozitív hőmérsékleti tartományon a rendszer rendezetlensége (entrópiája) nő a hőmérséklet emelkedésével.  $\square$

Adott  $(E_i)_{1 \leq i \leq n}$  energiaszintekhez és  $\beta$  paraméterhez definiáljuk a

$$R_{(E_i)_{1 \leq i \leq n}}^{(\beta)} = \left( \frac{e^{-\beta E_1}}{\sum_{i=1}^n e^{-\beta E_i}}, \dots, \frac{e^{-\beta E_n}}{\sum_{i=1}^n e^{-\beta E_i}} \right) \quad (1.75)$$

*Gibbs-féle valószínűségi eloszlást.*

A fenti példa gondolatmenete általánosabb esetben is végigkövethető. Tekintsük az  $X_n = \{0, \dots, n\}$  véges halmazt, és legyen  $(X_n, M, \Xi)$  az  $X_n$  alaphalmazú diszkrét eloszlásból álló statisztikai modell. (Lásd az (1.1.) példát). Legyenek továbbá  $k \leq n$  esetén  $F_1, \dots, F_k : X_n \rightarrow \mathbb{R}$  tetszőleges függvények. (Az  $F_i$  függvények az adott  $x \in X_n$  állapotok különböző fizikai paramétereit írják le.) Jelölje  $S$  az

$$p(x, \vartheta) = \exp \left( \sum_{i=1}^k \vartheta_i F_i(x) - \psi(\vartheta) \right) \quad (1.76)$$

alakú eloszlások halmazát. Ekkor  $S \subseteq M$  teljesül. Adott  $a \in \mathbb{R}^k$  vektorra legyen  $M_a$  azon  $M$ -beli eloszlások halmaza, melyre az  $F_i$  mennyiségek várható értéke  $a_i$ , vagyis

$$M_a = \left\{ q \in M \mid \forall i \in \{1, \dots, k\} : \sum_{x \in X_n} q(x, \eta) F_i(x) = a_i \right\}. \quad (1.77)$$

Tegyük fel továbbá, hogy az  $S \cap M_a$  halmaz nem üres és  $p \in S \cap M_a$ . Ekkor

$$\max_{q \in M_a} S(q) = S(p) \quad (1.78)$$

teljesül, vagyis adott várható értékű eloszlások közül az exponenciális családban lévőknek lesz maximális az entrópiája. A maximálisentrópia-elv értelmében ez azt jelenti, hogy a fizikai rendszer (legvalószínűbb egyensúlyi) állapota az  $S$  exponenciális családban van.

További, maximális entrópia elvének megfelelő példákat lehet hozni a fizikában szereplő  $(X, S, \Xi)$  statisztikai modellekre.

Diszkrét eloszlások esetén, az entrópia függvényből is származtatható a Fisher-féle információ.

**1.5. Példa.** Tekintsük az (1.1.) példában bevezetett diszkrét eloszlást. Az entrópia függvény ekkor

$$S : \Xi \rightarrow \mathbb{R} \quad (p_0, p_1, \dots, p_n) = - \sum_{i=0}^n p_i \log p_i . \quad (1.79)$$

Használjuk ki, hogy az entrópia csak a  $p_1, \dots, p_n$  értékektől függ, hiszen ekkor

$$p_0 = 1 - p_1 - \dots - p_n . \quad (1.80)$$

Az  $R^{n+1}$  azon elemeit, melyre  $x_0 + x_1 + \dots + x_n = 0$  teljesül jelölje  $X^{n+1}$ . Az  $\underline{x} \in X^{n+1}$  vektornál szintén csak az  $x_1, \dots, x_n$  paraméterekkel foglalkozunk.

Az entrópia  $k$ -adik parciális deriváltja ( $k = 1, \dots, n$ )

$$\begin{aligned} \partial_k S(p_1, \dots, p_n) &= \partial_k \left( (p_1 + \dots + p_n - 1) \log(1 - p_1 - \dots - p_n) - \sum_{i=1}^n p_i \log p_i \right) \\ &= \log(1 - p_1 - \dots - p_n) - \log p_k . \end{aligned} \quad (1.81)$$

Ez alapján az entrópia deriváltjára

$$dS : \Xi \rightarrow \text{Lin}(X^{n+1}, \mathbb{R}) \quad p \mapsto \left( \underline{x} \mapsto dS(p)(\underline{x}) \right) \quad (1.82)$$

$$dS(p_1, \dots, p_n)(x_0, \dots, x_n) = \sum_{i=1}^n x_i \left( \log(1 - p_1 - \dots - p_n) - \log p_i \right)$$

teljesül. Az entrópia másodrendű parciális deriváltja ( $k, l = 1, \dots, n$ )

$$\partial_l \partial_k S(p_1, \dots, p_n) = \partial_l \left( \log(1 - p_1 - \dots - p_n) - \log p_k \right) = \delta_{kl} \frac{1}{p_k} + \frac{1}{1 - p_1 - \dots - p_n} . \quad (1.83)$$

Ezek alapján az entrópia második deriváltját az

$$d^2 S : \Xi \rightarrow \text{Lin}(X^{n+1} \times X^{n+1}, \mathbb{R}) \quad p \mapsto \left( (\underline{x}, \underline{y}) \mapsto d^2 S(p)(\underline{x}, \underline{y}) \right) \quad (1.84)$$

$$d^2 S(p_1, \dots, p_n)((x_0, \dots, x_n), (y_0, \dots, y_n)) = \sum_{i,j=1}^n x_i y_j \left( \delta_{ij} \frac{1}{p_i} + \frac{1}{1 - p_1 - \dots - p_n} \right)$$

alakban lehet felírni. Vagyis az  $x_i y_j$  tag együttthatója éppen a Fisher-féle információs mátrix  $i, j$ -edik tagjának a  $(-1)$ -szerese. Tehát az entrópia második deriváltja minden  $p \in \Xi$  pontban azonosítható a Fisher-féle információs mátrix  $(-1)$ -szeresével.  $\square$

A továbbiakban néhány általános megjegyzést teszünk egy eloszlás Fisher-féle információjának és Shannon-entrópiájának a kapcsolatáról. Fontos azonban kiemelni a leglényegesebb különbséget, hogy a Fisher-féle információt statisztikai modellekre is értelmeztük, míg az entrópiát eloszlásokra.

1. Egy eloszlás Fisher-féle információja szigorúan pozitív mennyiség, míg entrópiája tetszőleges valós szám lehet. Az információ mértékéül szolgáló pozitív mennyiséget könnyebb értelmezni, mint a rendezetlenséget meghatározó entrópiát. (Diszkrét eloszlások esetén az entrópia is szigorúan pozitív mennyiség.)
2. A  $p$  egyváltozós sűrűségfüggvény esetén a Fisher-féle információ a

$$g = 4 \int_{\mathbb{R}} \left( \frac{d \sqrt{p(x)}}{dx} \right)^2 dx \quad (1.85)$$

alakot ölti. A  $q(x) = \sqrt{p(x)}$  mennyiséget Fisher *valószínűségi amplitúdófüggvénynek* nevezte el, és fontos szerepet tulajdonított neki. Az

$$\mathcal{L} = 4(q(x)')^2 \quad (1.86)$$

mennyiséget Lagrange-sűrűségként értelmezte, majd az elméleti fizikában megszokott Lagrange-formalizmust általánosította a fizikai információelméletre, bevezetve a potenciális energia információelméleti megfelelőjét. Érdeemes megemlíteni, hogy Fisher komplex valószínűségi amplitúdó függvényeket vizsgálva írta fel a

$$\mathcal{L}_m = C \nabla \psi \cdot \nabla \psi^* \quad (1.87)$$

típusú mozgási energiát tartalmazó Lagrange-sűrűséget, melyet vele egy évben, 1926-ban, Schrödinger a kvantummechanikai állapotot jellemző  $\psi$  függvényre vezetett be.

3. A maximális entrópia elvének megvan a Fisher-féle információval leírható megfelelője. Ezek szerint magárahagyott rendszernek csökken az információtartalma. Vagyis ha egy rendszer állapotának  $t$  időbeli változását valamilyen  $f_t$  sűrűségfüggvény írja le, és minden  $t$  időpillanatban  $g(f_t)$  jelöli az  $f_t$  sűrűségfüggvény Fisher-féle információját, akkor

$$\frac{d g(f_t)}{dt} \leq 0 \quad (1.88)$$

teljesül [33].

A Fisher-információ további és mélyebb fizikai alkalmazásait mutatja be a [34] könyv.

Végül említsük meg, hogy Fisher információelmélethez szorosan kapcsolódó statisztikai vizsgálódásai [30, 31, 32] évtizedekkel Shannon első munkái előtt kezdődtek.

### 1.3. Eloszlások távolsága

A statisztika és a valószínűségszámítás különböző tudományterületeken való alkalmazása során komoly problémát jelent a megfigyelt eloszlások közötti eltérés (vagy másképp fogalmazva, távolság) fogalmának a tisztázása. A távolságnak a releváns értelmezése függ a tudományterülettől. Például az antropológiában [92], a genetikában [72], a ökológiában [90] és a mintázatfelismerésben [11] más-más távolságfogalmakat értelmeznek az eloszlások között. Az eloszlások közötti távolság általában nem szimmetrikus függvény. Gyakorlati alkalmazások esetén ez érthető, hiszen egy rendszer állapotának a megváltoztatása lehet, hogy kevesebb vagy több fáradságba kerül, mint a visszaállítása. A matematikai megfelelője ezen távolságfogalmaknak a divergencia.

**1.9. Definíció.** Legyen  $(X, S, \Xi)$  statisztikai modell. A

$$D : S \times S \rightarrow \mathbb{R} \quad (p, q) \mapsto D(p, q) \quad (1.89)$$

leképezést *általánosított kontrasztfüggvénynek* (vagy *általánosított divergenciának*) nevezünk, ha minden  $p, q \in S$  esetén  $D(p, q) \geq 0$ , valamint  $D(p, q) = 0$  pontosan akkor teljesül, ha  $p = q$ .

A  $D$  általánosított divergencia távolságjellegű függvény, azonban nem követeljük meg tőle a szimmetrikusságot és a háromszög-egyenlőtlenség teljesülését.

Az információelméletben és a statisztikában többféle általánosított divergenciát alkalmaznak.

$$\text{Kullback–Liebler-féle [59]: } D_{\text{KL}}(p, q) = \int_X p(x) \log \frac{p(x)}{q(x)} \, dx \quad (1.90)$$

$$\text{Hellinger-féle [50]: } D_{\text{H}}(p, q) = \int_X (\sqrt{p(x)} - \sqrt{q(x)})^2 \, dx \quad (1.91)$$

$$\chi^2: \quad D_{\chi^2}(p, q) = \int_X p(x) \left( \left( \frac{p(x)}{q(x)} \right)^2 - 1 \right) \, dx \quad (1.92)$$

$$\alpha \in ]-1, 1[ \text{ számra: } D_{\alpha}(p, q) = \frac{4}{1 - \alpha^2} \left[ 1 - \int_X p(x)^{\frac{1-\alpha}{2}} q(x)^{\frac{1+\alpha}{2}} \, dx \right] \quad (1.93)$$

$$\text{Bhattacharyya-féle [15]: } D_{\text{B}}(p, q) = 1 - \int_X \sqrt{p(x)q(x)} \, dx \quad (1.94)$$

$$\text{Harmonikus: } D_{\text{Ha}}(p, q) = 1 - \int_X \frac{2p(x)q(x)}{p(x) + q(x)} \, dx \quad (1.95)$$

$$\text{Jeffreys-féle [54]: } D_{\text{J}}(p, q) = \int_X (p(x) - q(x)) \log \frac{p(x)}{q(x)} \, dx \quad (1.96)$$

$$\text{Háromszög [103]: } D_{\Delta}(p, q) = \int_X \frac{(p(x) - q(x))^2}{p(x) + q(x)} \, dx \quad (1.97)$$

$$\text{Lin–Wong-féle [65]: } D_{\text{LW}}(p, q) = \int_X p(x) \log \frac{2p(x)}{p(x) + q(x)} \, dx \quad (1.98)$$

Ezek a mennyiségek, az eloszlások közötti távolság fogalmának a különböző megközelítéseiből származtathatók. Például a  $D_{\text{KL}}(p, q)$  mennyiség azt fejezi ki, hogy mennyi információt nyerhetünk a  $p$  eloszlásból származó megfigyelésből ahhoz, hogy úgy döntsünk, hogy megfigyeléseink a  $p$  eloszlásból származnak és nem a  $q$  eloszlásból; azaz mennyi információ különbözteti meg a  $p$  eloszlást  $q$ -tól [60].

Az eloszlások között bevezetett eltérések nagy részét egy alkalmas függvény segítségével általánosan lehet kezelni. Legyen az  $f : \mathbb{R} \rightarrow \mathbb{R}$  függvény szigorúan konvex a pozitív számok halmazán és legyen  $f(1) = 0$ . Az  $(X, \mathcal{S}, \Xi)$  statisztikai modellen a

$$D_f(p, q) = \int_X p(x) f \left( \frac{q(x)}{p(x)} \right) \, dx \quad (1.99)$$

függvényt *Csiszár-féle  $f$ -divergenciának* [22, 23] nevezzük. Minden  $c \in \mathbb{R}$  paraméterre  $D_{f(u)} = D_{f(u)+c(u-1)}$  teljesül, továbbá bevezetve az  $f^\wedge(u) = uf(u^{-1})$  függvényt

$$D_f(p, q) = D_{f^\wedge}(q, p) . \quad (1.100)$$

Legyen  $\alpha \in \mathbb{R}$  és

$$f_\alpha : \mathbb{R} \rightarrow \mathbb{R} \quad x \mapsto \begin{cases} \frac{4}{1-\alpha^2} \left(1 - x^{\frac{1+\alpha}{2}}\right) & \text{ha } \alpha \neq \pm 1 \\ x \log x & \text{ha } \alpha = 1 \\ -\log x & \text{ha } \alpha = -1 \end{cases} \quad (1.101)$$

függvény. Ekkor  $D_{f_{-1}} = D_{\text{KL}}$ ,  $D_{f_0} = 2D_{\text{H}}$ , továbbá  $\alpha \neq \pm 1$  esetén  $D_{f_\alpha} = D_\alpha$  teljesül. Az  $f_1(x) = (x-1)^2$  és az  $f_2(x) = 1 - \sqrt{x}$  esetben  $D_{f_1} = D_{\chi^2}$  és  $D_{f_2} = D_{\text{H}}$  teljesül.

Az fent definiált divergenciák között számos egyenlőtlenség írható fel, például [102]

$$D_{\text{Ha}}(p, q) \geq \exp\left(-\frac{1}{2}D_{\text{J}}(p, q)\right) \quad (1.102)$$

$$D_{\text{Ha}}(p, q) \geq 1 - \frac{1}{4}D_{\text{J}}(p, q) \quad (1.103)$$

$$2D_{\Delta}(p, q) \leq D_{\text{J}}(p, q) \leq \frac{D_{\chi^2}(p, q) + D_{\chi^2}(q, p)}{2} \quad (1.104)$$

A  $D_f$  Csiszár-féle  $f$  divergenciának hasonló monotonitási tulajdonsága van, mint a Fisher-féle információnak, és a két változójában együttesen konvex.

**1.6. Tétel.** Legyen  $p, q : X \rightarrow \mathbb{R}$  két sűrűségfüggvény,  $Y$  halmaz és  $\kappa : X \times Y \rightarrow \mathbb{R}$  átmenetvalószínűség. Legyen  $\tilde{p}(y) = \int_X \kappa(y|x)p(x) dx$ ,  $\tilde{q}(y) = \int_X \kappa(y|x)q(x) dx$  és  $D_f$  Csiszár-féle  $f$ -divergencia. Ekkor

$$D_f(p, q) \geq D_f(\tilde{p}, \tilde{q}) \quad (1.105)$$

teljesül.

**1.7. Tétel.** Legyen  $p_1, p_2, q_1, q_2 : X \rightarrow \mathbb{R}$  sűrűségfüggvény,  $0 \leq \lambda \leq 1$  paraméter és  $D_f$  Csiszár-féle  $f$ -divergencia. Ekkor

$$D_f(\lambda p_1 + (1-\lambda)p_2, \lambda q_1 + (1-\lambda)q_2) \leq \lambda D_f(p_1, q_1) + (1-\lambda)D_f(p_2, q_2) \quad (1.106)$$

teljesül.

Az általánosított divergencia fogalma túl tág keretet jelent a differenciálgeometriai analízishez. Legyen  $D$  olyan általánosított divergencia, hogy minden  $p \in S$  esetén az  $y \mapsto D(p(x, \vartheta + y), p(x, \vartheta))$  függvény sorbafejthető  $y$ -szerint. Az általánosított divergencia definíciója alapján a sorfejtésében az első nullától különböző elem a négyzetes tag lesz

$$D(p(x, \vartheta + y), p(x, \vartheta)) = \frac{1}{2} \sum_{i,k=1}^n g_{ik}^{(D)}(p) y_i y_k + \frac{1}{6} \sum_{i,j,k=1}^n h_{ijk}^{(D)} y_i y_j y_k + o(\|y\|^3). \quad (1.107)$$

**1.10. Definíció.** Azt mondjuk, hogy az  $(X, S, \Xi)$  statisztikai modellen a  $D$  általánosított divergencia *divergencia* (vagy *kontrasztfüggvény*), ha minden  $p \in S$  esetén a  $D(p(x, \vartheta + y), p(x, \vartheta))$  függvény sorbafejthető  $y$ -szerint minden  $\vartheta \in \Xi$  paraméter mellett, és a sorfejtésben megjelenő  $g_{ik}^{(D)}$  mátrix pozitív definit. A  $D^*(p, q) = D(q, p)$  képlettel értelmezett  $D^*$  függvényt a  $D$  általánosított divergencia *duális általánosított divergenciájának* nevezzük.

A duális divergencia sorfejtése

$$D^*(p(x, \vartheta + y), p(x, \vartheta)) = \frac{1}{2} \sum_{i,k=1}^n g_{ik}^{(D)}(p) y_i y_k + \frac{1}{6} \sum_{i,j,k=1}^n h_{ijk}^{(D^*)} y_i y_j y_k + o(\|y\|^3). \quad (1.108)$$

A másodrendű tagok tehát azonosak a divergenciánál és a duálisánál, ezért általánosan igaz, hogy egy divergencia duálisa divergencia.

Tekintsük a  $D_f$  Csiszár-féle  $f$ -divergenciát. Ha a  $g$  szigorúan konvex,  $g(1) = 0$  feltételt teljesítő függvényre

$$g(x) + xg(x^{-1}) = f(x) + xf(x^{-1}) \quad (1.109)$$

teljesül, akkor az (1.107) és (1.108) egyenlet alapján, a  $D_g$  Csiszár-féle divergencia ugyan azt a másodrendű tagot adja, mint a  $D_f$  divergencia, vagyis  $g_{ik}^{D_g} = g_{ik}^{D_f}$ .

**1.6. Példa.** A (1.1.) példában bemutatott diszkrét eloszlások esetén a Csiszár-féle  $f$ -divergenciák sorfejtésében megjelenő  $g^{D_f}$  másodrendű tagot többféle képpen lehet jellemezni. Jelölje  $\partial_i^{(p)}$  illetve  $\partial_i^{(q)}$  a

$$D_f : S \times S \rightarrow \mathbb{R} \quad (p, q) \mapsto D_f(p, q) \quad (1.110)$$

függvény  $p_i$  és  $q_i$  változó szerinti parciális deriváltját ( $i=0, \dots, n$ ). Rövid számolással ellenőrizhetők az alábbi azonosságok  $i, j = 1, \dots, n$  esetén.

$$(\partial_i^{(p)} \partial_j^{(p)} D_f)(p, p) = f''(1) \left( \delta_{ij} \frac{1}{p_i} + \frac{1}{p_0} \right) \quad (1.111)$$

$$(\partial_i^{(p)} \partial_j^{(q)} D_f)(p, p) = -f''(1) \left( \delta_{ij} \frac{1}{p_i} + \frac{1}{p_0} \right)$$

$$(\partial_i^{(q)} \partial_j^{(q)} D_f)(p, p) = f''(1) \left( \delta_{ij} \frac{1}{p_i} + \frac{1}{p_0} \right)$$

Ezek alapján minden  $\underline{x}, \underline{y} \in X^{n+1}$  vektorra minden  $p \in S$  pontban

$$\begin{aligned} \langle \underline{x}, g^{D_f}(p) \underline{y} \rangle &= \sum_{i,j=1}^n x_i y_j g_{ij}^{D_f}(p) = \frac{\partial^2}{\partial t \partial s} D_f(p + t\underline{x} + s\underline{y}, p) \Big|_{t,s=0} = \\ &= - \frac{\partial^2}{\partial t \partial s} D_f(p + t\underline{x}, p + s\underline{y}) \Big|_{t,s=0} = \\ &= \frac{\partial^2}{\partial t \partial s} D_f(p, p + t\underline{x} + s\underline{y}) \Big|_{t,s=0} \end{aligned} \quad (1.112)$$

teljesül.

□

A divergenciákból származó  $g_{ik}^{(D)}$  mátrix általában számszorzó erejéig megegyezik a Fisher-információs mátrixszal.

**1.8. Tétel.** *Legyen  $(X, S, \Xi)$  statisztikai modell. Ekkor*

$$\begin{aligned} g^{(D_{KL})} &= g^{(F)} & g^{(D_H)} &= \frac{1}{2} g^{(F)} & g^{(D_{\chi^2})} &= 2g^{(F)} & (1.113) \\ g^{(D_\alpha)} &= g^{(F)} & g^{(D_B)} &= \frac{1}{4} g^{(F)} & g^{(D_{H_a})} &= \frac{1}{2} g^{(F)} \\ g^{(D_J)} &= 2g^{(F)} & g^{(D_\Delta)} &= g^{(F)} & g^{(D_{LW})} &= \frac{1}{4} g^{(F)} \\ g^{(D_f)} &= f''(1)g^{(F)} \end{aligned}$$

teljesül.

Ezek alapján az általánosított divergenciára adott (1.90–1.98) példák egyúttal divergenciák is. Bizonyos esetekben az eloszlások között a mérhetőség segítségével definiálhatunk távolságfogalmat.

**1.7. Példa.** Wooters-példája [113]: Tegyük fel, hogy van két valószínűségi eloszlásunk:  $(p_1, 1 - p_1)$  és  $(p_2, 1 - p_2)$ , ahol  $p_1 < p_2$ . Ha  $n$  mérési lehetőségünk van, akkor a mérés bizonytalanságát jól jellemzi a

$$\Delta p = \sqrt{\frac{p(1-p)}{n}} \quad (1.114)$$

paraméter, mely a tipikus fluktuációk nagysága. Ezek alapján azt mondjuk, hogy a  $(p_1, 1 - p_1)$  és  $(p_2, 1 - p_2)$  eloszlások  $n$ -méréssel megkülönböztethetők, ha a tipikus fluktuációk között nincs átfedés, vagyis ha

$$|p_1 - p_2| \geq \Delta p_1 + \Delta p_2 \quad (1.115)$$



teljesül. Jelölje  $k(n, p_1, p_2)$  azon  $(p_i, 1 - p_i)$  valószínűségi eloszlások számát, amelyekre  $p_1 < p_i < p_2$ ,  $p_i < p_{i+1}$  teljesül és  $(p_i, 1 - p_i)$   $n$ -méréssel megkülönböztethető a  $(p_{i+1}, 1 - p_{i+1})$  eloszlástól. Ekkor a  $(p_1, 1 - p_1)$  és a  $(p_2, 1 - p_2)$  eloszlások között a statisztikai távolság legyen

$$d(p_1, p_2) = \lim_{n \rightarrow \infty} \frac{k(n, p_1, p_2)}{\sqrt{n}}. \quad (1.116)$$

Az előző egyenlet alapján a távolság

$$d(p_1, p_2) = \int_{p_1}^{p_2} \frac{1}{\sqrt{p(1-p)}} dp = \arccos \left( \sqrt{p_1 p_2} + \sqrt{(1-p_1)(1-p_2)} \right). \quad (1.117)$$

□

## 1.4. Majorizáció diszkrét eloszlásoknál

Hardy, Littlewood és Pólya 1934-ben megjelent könyve az elsők között tárgyalja alaposan a majorizációt és a hozzá kapcsolódó egyenlőtlenségeket [43]. Részletesebb és általánosabb leírás található a majorizációról Marshall és Olkin 1979-ben megjelent könyvében [70]. A téma rövid, de átfogó tárgyalása megtalálható még Bhatia [14] könyvében és Ando [7] cikkében.

Ebben a fejezetben a véges halmazon értelmezett valószínűségi mértékekkel foglalkozunk. Tehát olyan  $(X, S, \Xi)$  statisztikai modelleket vizsgálunk, ahol  $X$  véges halmaz.

Adott  $(a_1, \dots, a_n) \in \mathbb{R}^n$  vektor esetén jelölje  $(a_1^\downarrow, \dots, a_n^\downarrow)$  az  $\{a_1, \dots, a_n\}$  számok nagyság szerinti csökkenő sorrendjét.

**1.11. Definíció.** Legyen  $p, q \in S$ , és legyenek  $(p_1^\downarrow, \dots, p_n^\downarrow)$ ,  $(q_1^\downarrow, \dots, q_n^\downarrow)$  a  $p$  és  $q$  csökkenő sorrendbe rendezett értékei. Azt mondjuk, hogy  $p$  *majorálja*  $q$ -t (vagy  $q$  *kevvertebb*  $p$ -nél), ha minden  $k = 1, \dots, n$  esetén

$$\sum_{i=1}^k q_i^\downarrow \leq \sum_{i=1}^k p_i^\downarrow \quad (1.118)$$

teljesül. Ezen reláció fennállását  $q \prec p$  a jelöli.

A  $q \prec p$  majorizációs reláció egyik fizikai interpretálása az, hogy  $q$  rendezetlenebb, kaotikusabb állapotot ír le mint  $p$ . A definícióból következik, hogy minden  $p \in S$  eloszlásra

$$\left( \frac{1}{n}, \dots, \frac{1}{n} \right) \prec (p_1, \dots, p_n) \prec (1, 0, \dots, 0) \quad (1.119)$$

teljesül.

Adott  $p \in S$  eloszlásból keverttebbet kapunk, ha az elemeiket közelítjük egymáshoz. Ha  $(p_1^\downarrow, \dots, p_n^\downarrow)$  jelöli  $p$  elemeit csökkenő sorrendben, akkor minden  $1 \leq i < j \leq n$  és  $t \in [0, 1]$  esetén

$$(p_1^\downarrow, \dots, tp_i^\downarrow + (1-t)p_j^\downarrow, \dots, (1-t)p_i^\downarrow + tp_j^\downarrow, \dots, p_n^\downarrow) \prec (p_1^\downarrow, \dots, p_i^\downarrow, \dots, p_j^\downarrow, \dots, p_n^\downarrow) \quad (1.120)$$

teljesül. A valószínűségek fenti keverését leíró

$$T_{ij}^{(t)}(p) = (p_1, \dots, tp_i + (1-t)p_j, \dots, (1-t)p_i + tp_j, \dots, p_n) \quad (1.121)$$

leképezést *T-transzformációnak* nevezzük.

Ha egy rendszert Gibbs-eloszlás jellemez, akkor a hőmérséklet emelkedésével kaotikusabb állapotot kapunk, ezt fogalmazza meg Wehrl tétele [111] bizonyos kvantumállapotokra, melyből azonnal adódik az alábbi egyszerűbb tétel.

**1.9. Tétel.** *Adott  $(E_i)_{1 \leq i \leq n}$  pozitív energiaszintek és  $\beta_1 < \beta_2$  paraméterek esetén*

$$R_{(E_i)}^{(\beta_1)} \prec R_{(E_i)}^{(\beta_2)} \quad (1.122)$$

teljesül, ahol  $R_{(E_i)}^{(\beta)}$  az (1.75) képlettel értelmezett Gibbs-féle eloszlás.

**1.12. Definíció.** Egy  $n \times n$ -es  $A$  mátrixot *kétszeresen sztochasztikusnak* nevezünk, ha minden eleme nemnegatív, továbbá

$$\sum_{i=1}^n a_{ij} = \sum_{j=1}^n a_{ij} = 1 \quad (1.123)$$

teljesül.

Minden  $n \times n$ -es kétszeresen sztochasztikus mátrix egyben véges halmazokon értelmezett átmenetvalószínűség is.

Adott  $p \in S$  eloszlás kétszeresen sztochasztikus mátrix általi képe szintén eloszlás lesz, és keverttebb lesz az eredetinel. Ennél azonban több is igaz.

**1.10. Tétel.** *Legyen  $p, q \in S$ . Ekkor az alábbiak ekvivalensek.*

1. *A  $p \prec q$  reláció teljesül.*
2. *Léteznek olyan  $p_0, \dots, p_k \in S$  eloszlások, melyekre  $p_0 = p, p_k = q$ , továbbá*

$$p_0 \prec \dots \prec p_k \quad (1.124)$$

teljesül, valamint minden  $i = 0, \dots, k-1$  esetén  $p_i$  a  $p_{i+1}$  eloszlás *T-transzformációjával* kapható meg.

3. Léteznek olyan  $p_0, \dots, p_m \in S$  eloszlások, melyekre  $p_0 = p$ ,  $p_m = q$ , továbbá

$$p_0 \prec \dots \prec p_m \tag{1.125}$$

teljesül, valamint minden  $i = 0, \dots, m - 1$  esetén léteznek  $(E_{i,j})_{1 \leq j \leq n}$  energiaszintek,  $\beta_{1,i}$  és  $\beta_{2,i}$  paraméterek, hogy  $p_i = R_{(E_{i,j})}^{(\beta_{1,i})}$  és  $p_{i+1} = R_{(E_{i,j})}^{(\beta_{2,i})}$  teljesül.

4. Van olyan  $A$  duplán sztochasztikus mátrix, hogy  $Aq = p$ .

Az  $x \mapsto -x \log x$  entrópia-függvény konkavitásából és a Jensen-egyenlőtlenségből adódik, hogy  $p \prec q$  esetén  $S(q) \leq S(p)$  teljesül. Tehát a kevertebb állapotnak nagyobb a rendezetlensége.



## 2. Fisher-metrika a klasszikus esetben

A fejezet célja a statisztikában alkalmazott differenciálgeometriai módszerek alapjainak a bemutatása. Az előző fejezet elején tett megkötések a vizsgálandó  $(X, S, \Xi)$  statisztikai modellre továbbra is érvényben maradnak, azonban a  $\Xi$  paraméterter újabb struktúrával bővül. Az alapvető újdonság a differenciálgeometriai megközelítésben, hogy a  $\Xi$  paraméterteret geometriai objektumnak (például differenciálgeometriának vagy Riemann-sokaságnak) tekinti, és mint a tapasztalat mutatja, a  $\Xi$  tér geometriai jellemzői (görbület, párhuzamos eltolás) statisztikai jelentéssel ruházhatók fel. Vagyis a differenciálgeometria mint hatékony eszköz szerepel ebben a megközelítésben.

A fejezetben főleg a diszkrét eloszlások geometriai tulajdonságait elemezzük, ugyanis ezt az elemzést lehet jól kiterjeszteni a nemkommutatív esetre. Az ehhez szükséges differenciálgeometriai fogalmakat vázlatosan áttekintjük. Azonban a fogalmakat nem a legáltalánosabb esetre definiáljuk, hanem megelégszünk azzal, hogy az általunk vizsgált esetekre alkalmazhatóak. Továbbá, ha a megszokott differenciálgeometriai definíciók több előkészítést igényelnek, akkor megelégszünk egy, az általunk vizsgált esetekben jól működő, a megszokottal ekvivalens definícióval. A differenciálgeometriai tételeket nem bizonyítjuk, és külön referenciát sem jelölünk meg, ha azok a bevezető jellegű differenciálgeometriai könyvekben megtalálhatóak. A számtalan általános differenciálgeometriai témájú könyv közül megemlítjük a magyar nyelvű [99, 100] valamint az angol nyelvű [49, 56, 63, 96] könyveket. A differenciálgeometria statisztikai alkalmazásainak néhány fejezete megtalálható a [2, 3, 10, 75] könyvekben.

Először a diszkrét eloszlások példáján mutatjuk be a differenciálgeometria fogalmait, majd az euklideszi terekben megszokott egyenes fogalmának az általánosítását, a geodetikust definiáljuk és elemezzük. Görbült térben egy pont körüli gömb térfogata általában különbözik a megszokott euklideszi térben számolt térfogattól. Ez az eltérés annál nagyobb, minél nagyobb a gömb sugara. Megvizsgáljuk, hogy a térfogat sugár szerinti sorfejtéséből hogyan nyerhetünk információt a térnek a középpontban való görbületéről és a görbülethez kapcsolódó egyéb mérőszámokról. Majd Cencov egyik tételét mutatjuk be, mely szerint létezik kitüntetett geometriai struktúra a diszkrét eloszlások terén. A nemkommutatív esetre előretekintve új módszerrel határozzuk meg a diszkrét eloszlások terének a görbületét. Az eddigiek alkalmazásaként megmutatjuk, hogy a normális eloszlások családja esetén hogyan használhatók ki és számolhatók a fejezetben definiált mennyiségek. Ennek az alkalmazásnak a szerepe a diszkrét eloszlások nemkommutatív általánosításánál derül ki. Majd megvizsgáljuk, hogy a paraméterter két pontját összekötő általánosított egyenes (geodetikus) mentén mért távolság milyen távolságfogalmat jelent a főbb eloszláscsaládok esetén.

## 2.1. A diszkrét eloszlások Riemann-geometriája

A fejezetben az (1.1.) példában bemutatott diszkrét eloszlás kevert paraméterezését elemezzük differenciálgeometriai szempontok alapján.

Legyen  $X_n = \{0, \dots, n\}$ , ekkor az  $X_n$  halmazon értelmezett sűrűségfüggvények  $n$  független paraméterrel adhatók meg

$$p(x, \vartheta_0, \dots, \vartheta_n) = \vartheta_i, \quad \text{ha } x = i, \quad 0 \leq i \leq n, \quad (2.1)$$

ahol  $\vartheta_0 + \dots + \vartheta_n = 1$ . A valódi  $n$ -változós eloszlások halmaza

$$\mathcal{P}_n = \left\{ p(x, \vartheta_0, \dots, \vartheta_n) \mid \forall i \in \{0, \dots, n\} : 0 < \vartheta_i < 1, \quad \sum_{i=0}^n \vartheta_i = 1 \right\}. \quad (2.2)$$

Ha  $\vartheta_1, \dots, \vartheta_n$ -et tekintjük függetleneknek, akkor  $\vartheta_0 = 1 - \sum_{k=1}^n \vartheta_k$ . A  $\mathcal{P}_n$  halmaz az  $(n+1)$  dimenziós téren egy hipersík pozitív ortánsba eső részével (egy szimplexszel) azonosítható.

A  $\mathcal{P}_n$  téren értelmezett függvényeket gyakran  $f(\vartheta_0, \dots, \vartheta_n)$  függvényként vagy  $f(\vartheta_1, \dots, \vartheta_n)$  függvényként írjuk, attól függően, hogy elvégeztük-e a  $\vartheta_0 = 1 - \sum_{k=1}^n \vartheta_k$  helyettesítést.

**2.1. Definíció.** Az  $(M, \mathcal{A})$  pár  $n$ -dimenziós differenciálható sokaság (sima sokaság vagy  $C^\infty$ -sokaság), ha

1. az  $M$  megszámlálható bázisú, Hausdorff-féle topologikus tér
2. az  $\mathcal{A}$  halmaz megszámlálható, és elemei olyan  $\phi_i : U_i \rightarrow V_i$  homeomorfizmusok, ahol  $U_i \subseteq M$  és  $V_i \subseteq \mathbb{R}^n$  nyílt halmazok,
3. minden  $\phi_i, \phi_j \in \mathcal{A}$  függvénypárra a

$$\phi_i \circ \phi_j^{-1} : \phi_j(U_i \cap U_j) \rightarrow \phi_i(U_i \cap U_j)$$

leképezés  $C^\infty$ -beli,

4. minden  $x \in M$  pont eleme valamelyik  $U_i$  halmaznak.

A  $\phi_i$  leképezéseket gyakran *lokális koordinátarendszereknek* (vagy *lokális koordinátázásoknak* vagy *lokális térképnek*) nevezzük. Legyen  $\phi$  egy lokális térkép. A  $\phi : U \rightarrow \mathbb{R}^n$  leképezést  $(x_1, \dots, x_n)$  alakba írhatjuk, ahol  $x_k = \text{pr}_k \circ \phi$ ,  $\text{pr}_k : \mathbb{R}^n \rightarrow \mathbb{R}$  pedig az  $i$ -edik kanonikus projekció.

Az  $(M, \mathcal{A})$  differenciálható sokaságot gyakran csak  $M$ -mel jelöljük, ha ez a megértést nem zavarja. A fentiek értelmében  $\mathcal{P}_n$  olyan  $n$ -dimenziós differenciálható sokaság, ahol az  $\mathcal{A}$  halmaznak csak a

$$\phi : \mathcal{P}_n \rightarrow \mathbb{R}^n \quad p(x, \vartheta_0, \dots, \vartheta_n) \mapsto (\vartheta_1, \dots, \vartheta_n) \quad (2.3)$$

leképezés az eleme.

**2.2. Definíció.** Legyen  $M$  differenciálható sokaság. Az  $f : M \rightarrow \mathbb{R}$  leképezésről azt mondjuk, hogy

- *differenciálható*, ha minden  $\phi$  lokális koordinátázásra  $f \circ \phi^{-1}$  differenciálható,
- *sima* (vagy  $C^\infty$ -beli), ha minden  $\phi$  lokális koordinátázásra az  $f \circ \phi^{-1}$  függvény  $C^\infty$ -beli.

**2.3. Definíció.** Legyen  $(M, \mathcal{A})$  és  $(N, \mathcal{B})$   $m$ - ill.  $n$ -dimenziós differenciálható sokaság. Az  $f : M \rightarrow N$  leképezés  $C^k$ -beli, ha minden  $(p, \phi_M)$  és  $(f(p), \phi_N)$  esetén a

$$\phi_N \circ f \circ \phi_M^{-1} : \mathcal{R}^n \rightarrow \mathcal{R}^m \quad (2.4)$$

leképezés  $C^k$ -beli, ahol  $p \in M$  és  $\phi_M \in \mathcal{A}$  a  $p$  pont egy környezetének a térképezése és  $\phi_N \in \mathcal{B}$  az  $f(p)$  pont egy környezetének a térképezése. Az  $f : M \rightarrow N$   $C^k$ -leképezés  $C^k$ -diffeomorfizmus, ha  $f$  és  $f^{-1}$  is  $C^k$  leképezések.

**2.1. Tétel.** Ha  $f : M \rightarrow N$  legalább  $C^1$ -diffeomorfizmus, akkor létezik  $\tilde{f} : M \rightarrow N$   $C^\infty$ -diffeomorfizmus.

A továbbiakban az  $f$  leképezést *diffeomorfizmusnak* hívjuk, ha legalább  $C^1$ -diffeomorfizmus.

Legyen  $M$  differenciálható sokaság, és  $p \in M$ . Jelölje  $\mathcal{F}_p$  azon  $M \rightarrow \mathbb{R}$  differenciálható függvények halmazát, melyek a  $p$  pont egy környezetében értelmezve vannak és jelölje  $\mathcal{F}(M)$  az egész  $M$  téren értelmezett differenciálható függvények halmazát. Az  $f_1, f_2 \in \mathcal{F}_p$  függvények  $f_1 \simeq f_2$  relációban állnak egymással, ha létezik a  $p$  pontnak olyan  $U$  környezete, amin az  $f_1$  és  $f_2$  függvények megegyeznek. Az  $\simeq$  reláció ekvivalenciareláció. Az  $(\mathcal{F}_p / \simeq)$  halmazt  $p$ -beli *függvénycsírának* nevezzük.

**2.4. Definíció.** Legyen  $M$  differenciálható sokaság és  $p \in M$ . A

$$D : (\mathcal{F}_p / \simeq) \rightarrow \mathbb{R} \quad (2.5)$$

leképezés *deriváció*, ha minden  $a$  és  $b$  valós számra és  $f, g \in \mathcal{F}_p / \simeq$  függvényre:

$$D(af + bg) = aD(f) + bD(g) \quad D(fg) = f(p)D(g) + D(f)g(p) \quad (2.6)$$

teljesül. A  $p$ -beli függvénycsírán értelmezett derivációk halmaza a *sokaság  $p$ -beli érintőtere*, jele  $T_p M$ .

A  $T_pM$  tér  $n$ -dimenziós vektortér. Az érintőterek halmazának

$$TM = \bigcup_{p \in M} T_pM \quad (2.7)$$

egyesítését *érintőnyalábnak* hívjuk.

Legyen  $\gamma : ] - \varepsilon, \varepsilon[ \rightarrow M$  olyan sima görbe, melyre  $\gamma(0) = p$  teljesül. Ekkor minden  $f : U_p \rightarrow \mathbb{R}$  differenciálható függvényre (ahol  $U_p$  a  $p$  pont egy nyílt környezete) értelmezett a

$$D_\gamma(f) = \left. \frac{df(\gamma(t))}{dt} \right|_{t=0} \quad (2.8)$$

leképezés (mely a  $\gamma$  görbe mentén vett deriválás) deriváció, azaz  $D_\gamma \in T_pM$ . Másrészt az érintőtér minden eleme előáll ilyen alakban.

Legyen  $p(x, \vartheta_0, \dots, \vartheta_n) \in \mathcal{P}_n$ . Adott  $(a_0, \dots, a_n)$  paraméterek esetén, melyekre  $a_0 + \dots + a_n = 0$  teljesül, létezik olyan  $\varepsilon > 0$ , hogy

$$\gamma : ] - \varepsilon, \varepsilon[ \rightarrow \mathcal{P}_n \quad t \mapsto p(x, \vartheta_0 + ta_0, \dots, \vartheta_n + ta_n) \quad (2.9)$$

jól definiált. A  $\gamma$  görbe menti deriválás ( $D_\gamma$ ) deriváció. Az érintőtér azonosítható azon  $(a_0, \dots, a_n)$  elemek halmazával melyekre  $a_0 + \dots + a_n = 0$  teljesül.

Legyen  $p \in M$  és tekintsük a  $p$  körüli  $\phi : U \rightarrow \mathbb{R}^n$  lokális koordinátarendszert. Ekkor minden  $f \in \mathcal{F}_p$  esetén legyen

$$\frac{\partial f}{\partial x_i} = \partial_i(f \circ \phi^{-1}), \quad (2.10)$$

ahol  $\partial_i$  jelöli az  $i$ -edik változó szerinti parciális deriváltat. Ekkor  $\frac{\partial}{\partial x_i}$  deriváció, melynek rövid jele  $\partial_i^{(\phi)}$ , vagy ha a megértést nem zavarja csak  $\partial_i$ . (Figyelem:  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  függvény esetén  $\partial_i f$  az  $i$ -edik változó szerinti parciális derivált, azonban  $f : M \rightarrow \mathbb{R}$  függvény esetén  $\partial_i f$  a fenti képlettel értelmezett deriváció!)

Visszatérve ismét a diszkrét eloszlásokra, adott  $(a_0, \dots, a_n)$  paraméterek esetén a (2.8) képlettel értelmezett görbe menti deriválásra

$$D_\gamma(f) = \frac{df(p(x, \vartheta_0 + ta_0, \dots, \vartheta_n + ta_n))}{dt} = (a_0 \partial_0 + \dots + a_n \partial_n) f \quad (2.11)$$

teljesül. Vagyis a  $\partial_0, \dots, \partial_n$  vektorok nem lineárisan függetlenek.

**2.5. Definíció.** Legyen  $M$  differenciálható sokaság. Egy

$$X : M \rightarrow TM \quad p \mapsto X(p) \quad (2.12)$$

leképezés  $C^\infty$ -vektormező (vagy röviden *vektormező*), ha



1. minden  $p \in \text{Dom}(X)$  esetén  $X(p) \in T_p M$ ,
2. a  $\text{Dom}(X)$  halmaz nyílt,
3. minden  $p \in \text{Dom}(X)$  és  $f \in \mathcal{F}_p$   $C^\infty$ -beli függvény esetén az

$$Xf : \text{Dom}(X) \cap \text{Dom}(f) \rightarrow \mathbb{R} \quad p \mapsto X(p)f \quad (2.13)$$

függvény  $C^\infty$ -beli.

Jelölje  $\mathcal{X}(M)$  az egész  $M$ -en értelmezett  $C^\infty$ -beli vektormezők halmazát.

Legyen  $X, Y \in \mathcal{X}(M)$  és  $f \in \mathcal{F}(M)$ . Ekkor az

$$Yf : M \rightarrow \mathbb{R} \quad p \mapsto Y(p)(f) \quad (2.14)$$

függvényre alkalmazható  $X$ :

$$X(Yf) : M \rightarrow \mathbb{R} \quad p \mapsto X(p)(Yf). \quad (2.15)$$

Ezáltal értelmezhető a vektormezők szorzata

$$XY : M \rightarrow \text{Lin}(\mathcal{F}(M), \mathbb{R}) \quad p \mapsto (f \mapsto X(p)(Yf)), \quad (2.16)$$

ami általában nem vektormező. A *vektormezők*  $[X, Y] = XY - YX$  *kommutátora* már vektormező lesz. Vagyis a  $\mathcal{X}(M)$  halmazon értelmezhető természetes módon a

$$[\cdot, \cdot] : \mathcal{X}(M) \times \mathcal{X}(M) \rightarrow \mathcal{X}(M) \quad (X, Y) \mapsto [X, Y] \quad (2.17)$$

leképezés.

Adott  $X \in \mathcal{X}(M)$  és  $f \in \mathcal{F}(M)$  esetén értelmezhető a

$$fX : M \rightarrow TM \quad p \mapsto f(p)X(p) \quad (2.18)$$

vektormező.

A Riemann-sokaság olyan tér, ahol a sokaság minden pontjának az érintőterében adott egy skalárszorzat.

**2.6. Definíció.** Legyen  $M$  differenciálható sokaság. A

$$g : M \rightarrow \text{Lin}(TM \times TM, \mathbb{R}) \quad (2.19)$$

leképezés *Riemann-metrika*, ha

1. minden  $p \in M$  esetén  $g_p : T_p M \times T_p M \rightarrow \mathbb{R}$  szimmetrikus, pozitív definit, bilineáris, nem degenerált leképezés (skalárszorzás),

2. minden  $X \in \mathcal{X}(M)$  esetén a

$$g(X, X) : M \rightarrow \mathbb{R} \quad p \mapsto g_p(X_p, X_p) \quad (2.20)$$

leképezés  $C^\infty$ -beli.

Az  $(M, g)$  párt *Riemann-sokaságnak* (vagy *Riemann-geometriának*) nevezzük, ha  $M$  differenciálható sokaság és  $g$  Riemann-metrika az  $M$  sokaságon.

A  $g$  leképezést gyakran csak metrikának, metrikus tenzornak nevezzük. A  $g_p$  metrika függ a  $p \in M$  ponttól, azonban, ha ez nem okoz félreértést egyszerűen csak  $g$ -t írunk  $g_p$  helyett. A  $p \in M$  pont körüli  $\phi$  lokális koordinátarendszerrel megadhatjuk a  $T_pM$  tér  $(\partial_1, \dots, \partial_n)$  bázisát. A  $g$  metrikát ebben a pontban a

$$g_{ij} = g(\partial_i, \partial_j) \quad (2.21)$$

képlettel értelmezett  $g_{ij}$  mátrix egyértelműen meghatározza.

A (1.8.) tétel alapján a főbb divergenciák infinitezimális formái mind a Fisher-mátrix számszorosát generálják. Az (1.1.) tétel alapján a  $\mathcal{P}_n$  sokaságon a Fisher-mátrix Riemann-metrikát határoz meg. Ennek  $g_{ij}$  metrikus tenzora

$$g_{ij}(\underline{\vartheta}) = \sum_{x=0}^n \frac{1}{p(x, \underline{\vartheta})} \frac{\partial p(x, \underline{\vartheta})}{\partial \vartheta_i} \frac{\partial p(x, \underline{\vartheta})}{\partial \vartheta_j} = \delta_{ij} \frac{1}{\vartheta_i} + \frac{1}{1 - \sum_{k=1}^n \vartheta_k}, \quad (2.22)$$

ahol  $\underline{\vartheta} = (\vartheta_1, \dots, \vartheta_n)$ . A  $\mathcal{P}_n$  téren a  $g$  Riemann-metrika mindig ezt a *Fisher-metrikát* jelöli a későbbiekben. Először Rao javasolta 1945-ben a Fisher-féle információs mátrix Riemann-metrikaként való használatát [91].

A továbbiakban a rövidebb írásmód kedvéért a bevezetendő  $\underline{\vartheta}$ -tól függő függvényeknél nem írjuk ki a  $\underline{\vartheta}$  argumentumot. A  $g_{ij}$  mátrix inverzét jelölje  $g^{ij}$ , ekkor

$$g^{ij} = -\vartheta_i \vartheta_j + \delta_{ij} \vartheta_i \quad (2.23)$$

teljesül.

A számolások során alkalmazzuk az *Einstein-féle írásmódot*, vagyis a képletekben az összeadandó tagokat minden olyan indexre összegezni kell (1-től  $n$ -ig), mely fent és lent is megjelenik, pl.

$$g_{ij} g^{jk} = \sum_{j=1}^n g_{ij} g^{jk}, \quad a_{i.k}^{j.} b^{lkm} c_{jm} = \sum_{j,k,m=1}^n a_{i.k}^{j.} b^{lkm} c_{jm} \quad g_i^i = \sum_{i=1}^n g_i^i. \quad (2.24)$$

**2.7. Definíció.** Legyen  $M$  differenciálható sokaság. A

$$\nabla : \mathcal{X}(M) \times \mathcal{X}(M) \rightarrow \mathcal{X}(M) \quad (X, Y) \mapsto \nabla_X Y \quad (2.25)$$

leképezés *kovariáns deriválás* (vagy *konnexió*), ha

1. minden  $X, Y, Z \in \mathcal{X}(M)$  vektormezőre

$$\nabla_{X+Y}Z = \nabla_X Z + \nabla_Y Z, \quad \nabla_X(Y + Z) = \nabla_X Y + \nabla_X Z$$

teljesül (mindkét változóban lineáris),

2. minden  $X, Y \in \mathcal{X}(M)$  vektormezőre és  $f \in \mathcal{F}(M)$  függvényre

$$\nabla_{fX}Y = f\nabla_X Y, \quad \nabla_X(fY) = (Xf)Y + f\nabla_X Y$$

teljesül (az első változóban  $\mathcal{F}(M)$ -lineáris).

**2.2. Tétel.** *A  $\nabla_X Y$  vektormező  $p$  pontbeli értékéhez elég az  $X(p)$  vektort ismerni.*

Ezért adott  $p \in M$  és  $v \in T_p M$  esetén értelmezhető a

$$\nabla_v : \mathcal{X}(M) \rightarrow T_p M \quad X \mapsto (\nabla_v X)(p) \quad (2.26)$$

képlettel a  $v$  vektor szerinti kovariáns deriválás.

A  $p \in M$  pont körüli  $\phi$  lokális térképen a  $\nabla$  kovariáns deriválás értelmezi a

$$\Gamma_{ijk} = g(\nabla_{\partial_i} \partial_j, \partial_k) \quad (2.27)$$

képlettel a  $\Gamma_{ijk}$  elsőfajú Christoffel-szimbólumokat, valamint a

$$\Gamma_{ij}^{\cdot k} \partial_k = \nabla_{\partial_i} \partial_j \quad (2.28)$$

képlettel a  $\Gamma_{ij}^{\cdot k}$  másodfajú Christoffel-szimbólumokat. Ezek között az

$$\Gamma_{ij}^{\cdot k} = \Gamma_{ijl} g^{lk}, \quad \Gamma_{ijk} = \Gamma_{ij}^{\cdot l} g_{lk} \quad (2.29)$$

összefüggések teljesülnek. Ezek alapján a  $\nabla$  kovariáns deriválás egyértelműen jellemezhető az első- vagy a másodfajú Christoffel-szimbólumokkal.

**2.8. Definíció.** Az  $(M, \nabla)$  párt, ahol  $M$  differenciálható sokaság és  $\nabla$  kovariáns deriválás, *differenciálgeometriának* (vagy *affinösszefüggő sokaságnak*) nevezzük.

**2.9. Definíció.** Az  $(M, \nabla)$  differenciálgeometriát *torziómentesnek* nevezzük, ha minden  $1 \leq i, j, k \leq n$  indexre  $\Gamma_{ij}^{\cdot k} = \Gamma_{ji}^{\cdot k}$  teljesül.

**2.10. Definíció.** Az  $(M, g)$  Riemann-sokaságon értelmezett  $\nabla$  kovariáns deriválást *Riemann-konnexiónak* nevezzük, ha minden  $X, Y, Z \in \mathcal{X}(M)$  vektormezőre

$$Xg(Y, Z) = g(\nabla_X Y, Z) + g(Y, \nabla_X Z) \quad (2.30)$$

teljesül.

**2.11. Definíció.** Az  $(M, g)$  Riemann-sokaságon értelmezett  $\nabla$  kovariáns deriválást *Levi-Civita-féle konnaxiónak* nevezzük, ha  $\nabla$  torziómentes és Riemann-konnaxió.

**2.3. Tétel.** *Adott  $(M, g)$  Riemann sokaságon pontosan egy  $\nabla$  Levi-Civita-féle kovariáns deriválás létezik, mely a*

$$\Gamma_{ij}^{..m} = g^{km} \frac{1}{2} (\partial_i g_{jk} + \partial_j g_{ik} - \partial_k g_{ij}). \quad (2.31)$$

*másodfajú Christoffel-szimbólummal adható meg.*

**2.12. Definíció.** Legyen  $-1 \leq \alpha \leq 1$ , ekkor a  $\mathcal{P}_n$  téren a

$$\Gamma_{ijk}^{(\alpha)} = \sum_{l=0}^n p(l, \underline{\vartheta}) \left( \partial_i \partial_j (\log p(l, \underline{\vartheta})) + \frac{1-\alpha}{2} (\partial_i \log p(l, \underline{\vartheta})) (\partial_j \log p(l, \underline{\vartheta})) \right) (\partial_k \log p(l, \underline{\vartheta})) \quad (2.32)$$

elsőfajú Christoffel-szimbólum által meghatározott  $\nabla^{(\alpha)}$  kovariáns deriválást  $\alpha$ -konnaxiónak (vagy  $\alpha$ -kovariáns deriválásnak) hívjuk.

Az  $\alpha$ -konnaxiókat Cencov vezette be, főbb tulajdonságait 1982-ben publikálta [19].

**2.4. Tétel.** *A  $(\mathcal{P}_n, g)$  téren, ahol  $g$  a Fisher-metrika, a  $\nabla^{(\alpha)}$   $\alpha$ -kovariáns deriválás a*

$$\Gamma_{ij}^{(\alpha)..k} = \frac{1+\alpha}{2} \left( \frac{\vartheta_k}{1 - \sum_{j=1}^n \vartheta_j} + \delta_{ij} \frac{\vartheta_k}{\vartheta_i} - \delta_{ij} \delta_{jk} \frac{1}{\vartheta_j} \right) \quad (2.33)$$

*másodfajú Christoffel-szimbólummal jellemezhető.*

**2.5. Tétel.** *A 0-kovariáns deriválás a Levi-Civita-féle kovariáns deriválás, vagyis  $\nabla^{(0)} = \nabla$ .*

**2.13. Definíció.** Legyen  $(M, \nabla)$  differenciálgeometria. A

$$R : \mathcal{X}(M) \times \mathcal{X}(M) \times \mathcal{X}(M) \rightarrow \mathcal{X}(M) \quad (X, Y, Z) \mapsto R(X, Y)Z \quad (2.34)$$

$$R(X, Y)Z = \nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z - \nabla_{[X, Y]} Z$$

leképezést *görbületi tenzornak* nevezzük. Azt mondjuk, hogy az  $(M, \nabla)$  differenciálgeometria *lapos*, ha minden  $X, Y, Z \in \mathcal{X}(M)$  esetén  $R(X, Y)Z = 0$  teljesül.

Statisztikai sokaságon a görbület fogalmát először Efron [28] kezdte vizsgálni 1975-ben. A statisztikus fizikán belül 1995-ben Ruppeiner figyelt fel a statisztikus sokaság görbületének fontosabb tulajdonságaira, például megmutatta, hogy a másodrendű fázisátalakulások esetén a kritikus pont közelében a görbület divergál, és a divergenciájának az exponense arányos bizonyos korrelációs függvények exponensével [93]. Az Ising modellek szimulációjánál fontos paraméter a spinlácok hossza. Egyik fő kérdés a szimulációk során, hogy a választott hossz, mennyire közelíti jól a „végtelent”. Ennek vizsgálatára 1999-ben Brody és Ritz a (Fisher-Rao metrikával vett) állapotter skalárgörbületét javasolta figyelendő paraméternek; megmutatták, hogy bizonyos modelleken az állapotter geometriája alapvetően más képet mutat kis láncméret ( $\leq 30$ ), mint nagy láncméret ( $\geq 45$ ) esetén [18].

Adott  $p \in M$  pont körüli lokális  $\phi$  térképen az  $R$  görbületi tenzort egyértelműen jellemzik a  $(\partial_1, \dots, \partial_n)$  bázison felvett értékei

$$R(\partial_i, \partial_j)\partial_k = R_{ijk}^{\dots l}\partial_l, \quad (2.35)$$

vagy az

$$g(R(\partial_i, \partial_j)\partial_k, \partial_l) = R_{ijkl} \quad (2.36)$$

mennyiségek, melyek között az

$$R_{ijkl} = R_{ijk}^{\dots m}g_{ml}, \quad R_{ijk}^{\dots l} = R_{ijkm}g^{ml} \quad (2.37)$$

összefüggések állnak fenn. Minden  $1 \leq i, j, k, l \leq n$  indexre a

$$R_{ijkl} = -R_{jikl}, \quad R_{ijkl} = -R_{ijlk}, \quad R_{ijkl} = R_{klij} \quad (2.38)$$

szimmetriatulajdonságok teljesülnek.

A másodfajú Christoffel-szimbólum segítségével kapjuk meg a Riemann-tenzor elemeit

$$R_{ijk}^{\dots l} = \partial_i\Gamma_{jk}^{\dots l} - \partial_j\Gamma_{ik}^{\dots l} + \Gamma_{jk}^{\dots m}\Gamma_{im}^{\dots l} - \Gamma_{ik}^{\dots m}\Gamma_{jm}^{\dots l}. \quad (2.39)$$

**2.6. Tétel.**  $A$   $(\mathcal{P}_n, \nabla^{(-1)})$  tér lapos.

**2.14. Definíció.** Legyen  $(M, \nabla)$  differenciálgeometria, melynek görbületi tenzora  $R$ . A

$$\text{Ric} : \mathcal{X}(M) \times \mathcal{X}(M) \rightarrow \mathcal{F}(M) \quad (X, Y) \mapsto \text{Tr}(Z \mapsto R(Z, X)Y) \quad (2.40)$$

leképezés a *Ricci-féle görbületi tenzor* (*Ricci-tenzor* vagy *Ricci-görbület*).

A  $p \in M$  pont körüli  $\phi$  lokális térképen a Ricci-tenzor egyértelműen megadható a

$$\text{Ric}_{ij} = \text{Ric}(\partial_i, \partial_j) \quad (2.41)$$

képlettel értelmezett  $\text{Ric}_{ij}$  mennyiségekkel. A Ricci-tenzor meghatározható a görbületi tenzorból:

$$\text{Ric}_{jk} = R_{ijk}^{\dots i}. \quad (2.42)$$

A  $(\mathcal{P}_n, g, \nabla^{(\alpha)})$  tér Ricci-tenzorának a meghatározásához vezessük be a

$$C = \left(1 - \sum_{i=1}^n \vartheta_i\right)^{-1} \quad (2.43)$$

jelölést. A Ricci-tenzorhoz az  $R_{ijk}^{(\alpha)\dots i}$  alakú elemeket kell kiszámolni, ezek a következők lesznek a fenti képlet alapján:

$$\partial_i \Gamma_{jk}^{(\alpha)\dots i} = \frac{1+\alpha}{2} \left( C + \vartheta_i C^2 + \delta_{jk} \frac{1}{\vartheta_j} \left(1 - \delta_{ij} + \delta_{ij} \frac{1}{\vartheta_j}\right) \right) \quad (2.44)$$

$$\partial_j \Gamma_{ik}^{(\alpha)\dots i} = \frac{1+\alpha}{2} \left( C + \vartheta_i C^2 + \delta_{ik} \delta_{ij} \frac{1}{\vartheta_i^2} \right)$$

$$\Gamma_{jk}^{(\alpha)\dots m} \Gamma_{im}^{(\alpha)\dots i} = \left( \frac{1+\alpha}{2} \right)^2 \left( \vartheta_i C^2 - C - \delta_{jk} \frac{1}{\vartheta_j} \left(1 + \delta_{ik} - \delta_{ik} \frac{1}{\vartheta_i}\right) \right)$$

$$\Gamma_{ik}^{(\alpha)\dots m} \Gamma_{jm}^{(\alpha)\dots i} = \left( \frac{1+\alpha}{2} \right)^2 \left( \vartheta_i C^2 - \delta_{ij} C - \delta_{ik} \frac{1}{\vartheta_i} \left( \delta_{kj} + \delta_{ij} - \delta_{kj} \frac{1}{\vartheta_i} \right) \right).$$

**2.7. Tétel.** A  $(\mathcal{P}_n, g, \nabla^{(\alpha)})$  tér Ricci tenzora

$$\text{Ric}_{jk}^{(\alpha)} = R_{ijk}^{(\alpha)\dots i} = \left( \frac{1+\alpha}{2} \right)^2 (n-1) \left( C + \frac{1-\alpha}{1+\alpha} \delta_{jk} \frac{1}{\vartheta_j} \right). \quad (2.45)$$

Legyen  $(M, g)$  Riemann-geometria és  $\text{Ric}$  a  $(M, \nabla)$  differenciálgeometria Ricci tenzora. Adott  $Y$  vektormező esetén a

$$\tilde{\text{Ric}}_Y : M \rightarrow \text{Lin}(\mathcal{X}(M), \mathbb{R}) \quad p \mapsto \text{Ric}(X, Y)(p) \quad (2.46)$$

leképezés egyértelműen meghatároz egy  $\tilde{Y}$  vektormezőt, melyre

$$\tilde{\text{Ric}}_Y(X)(p) = g_p(X, \tilde{Y}) \quad (2.47)$$

teljesül minden  $X \in \mathcal{X}(M)$  esetén. A

$$\text{Scal} : M \rightarrow \mathbb{R} \quad p \mapsto \text{Tr}(Y \rightarrow \tilde{Y}) \quad (2.48)$$

függvény az  $(M, g)$  Riemann-geometria  $\nabla$  kovariáns deriválásához tartozó *skalárgörbülete*.

A skalárgörbület a Ricci tenzor nyoma

$$\text{Scal} = \text{Ric}_{ij} g^{ij}. \quad (2.49)$$

**2.8. Tétel.** *A  $(\mathcal{P}_n, g, \nabla^{(\alpha)})$  tér skalárgörbülete*

$$\text{Scal}(\vartheta_1, \dots, \vartheta_n) = \frac{1 + \alpha}{4}(n - 1) \left( (1 - \alpha)n + 2\alpha \sum_{k=1}^n \vartheta_k \right). \quad (2.50)$$

A (2.5.) tétel értelmében mondhatjuk, hogy a  $(\mathcal{P}_n, g, \nabla^{(\alpha)})$  tér skalárgörbülete pontosan akkor állandó, ha  $\nabla^{(\alpha)}$  a Levi-Civita-féle kovariáns deriválás, vagyis ha  $\alpha = 0$ .

**2.15. Definíció.** Az  $(M, g)$  Riemann sokaságon egy  $\gamma : \mathbb{R} \rightarrow M$  sima függvény  $[a, b] \subseteq \mathbb{R}$  intervallumhoz tartozó ívének a *hossza*

$$l_\gamma(a, b) = \int_a^b \sqrt{g(\dot{\gamma}(t), \dot{\gamma}(t))} dt. \quad (2.51)$$

Az ívhossz a görbe átparaméterezésével szemben invariáns. A  $\mathcal{P}_n$  sokaság két pontja között az ívhossz segítségével definiálhatunk távolságot. Ez a távolság azonban függ a pontokat összekötő úttól.

A  $\mathcal{P}_1$  sokaság minden pontja egyetlen  $p$  paraméterrel jellemezhető. A  $p_1$  és  $p_2$  pontok közötti távolság meghatározásához tekintsük a  $\gamma(t) = t$  görbét. Ekkor

$$l(p_1, p_2) = \int_{p_1}^{p_2} \frac{1}{\sqrt{t(1-t)}} dt, \quad (2.52)$$

mely megegyezik a (1.117) képlettel.

Legyen  $(M, g)$  Riemann-sokaság,  $p \in M$  olyan pont, mely a  $\phi_1$  és  $\phi_2$  lokális koordinátarendszerek értelmezési tartományába esik, továbbá legyen  $U = \text{Dom}(\phi_1) \cap \text{Dom}(\phi_2)$ . A lokális koordinátarendszerekkel a  $T_p M$  tér két különböző bázisát adhatjuk meg:  $(\partial_1^{(1)}, \dots, \partial_n^{(1)})$ -et a  $\phi_1$  segítségével, valamint  $(\partial_1^{(2)}, \dots, \partial_n^{(2)})$ -et  $\phi_2$ -vel. A két bázisban a  $g$  Riemann-metrika a  $g_{ij}^{(1)}$  és a  $g_{ij}^{(2)}$  metrikus tenzort adja. Tekintsünk egy  $f : U \rightarrow \mathbb{R}$  sima függvényt. Ekkor

$$\int_{\phi_1(U)} \rho_1 \cdot (f \circ \phi_1^{-1}) = \int_{\phi_2(U)} \rho_2 \cdot (f \circ \phi_2^{-1}) \quad (2.53)$$

teljesül, ahol  $\rho_1 = \sqrt{\det g_{ij}^{(1)}}$  és  $\rho_2 = \sqrt{\det g_{ij}^{(2)}}$ . E tulajdonság miatt a  $\rho = \sqrt{\det g}$  függvényt *invariáns térfogati sűrűségfüggvénynek* nevezik. Ennek a segítségével értelmezhető a sokaság  $U \subseteq \text{Dom}(\phi)$  részének a *térfogata*:

$$V(U) = \int_{\phi(U)} \rho. \quad (2.54)$$

Nagyobb  $U \subseteq M$  tartományok térfogatát a tartomány több kisebb diszjunkt részre való felosztásával lehet kiszámolni.

A  $\mathcal{P}_n$  sokaságon

$$\rho = \left[ \left( 1 - \sum_{k=1}^n \vartheta_k \right) \prod_{k=1}^n \vartheta_k \right]^{-\frac{1}{2}}. \quad (2.55)$$

Ezek alapján meghatározható a  $\mathcal{P}_n$  tér térfogata:

$$V(\mathcal{P}_n) = \int_{0 < \vartheta_i < 1, \sum_{i=1}^n \vartheta_i < 1} \dots \int \left[ \left( 1 - \sum_{k=1}^n \vartheta_k \right) \prod_{k=1}^n \vartheta_k \right]^{-\frac{1}{2}} d\vartheta_1 \dots d\vartheta_n. \quad (2.56)$$

Ezt az  $a_i := \sqrt{\vartheta_i}$  és  $a_0 := \sqrt{1 - \sum_{k=1}^n \vartheta_k}$  helyettesítéssel vissza lehet vezetni arra, hogy egy konstans függvényt az  $(n+1)$  dimenziós térben lévő egységgömb pozitív térrészbe eső részén integrálunk.

**2.9. Tétel.** *Az  $\mathcal{P}_n$  sokaság térfogata megegyezik az  $n$ -dimenziós gömb felszínével*

$$V(\mathcal{P}_n) = \frac{2\pi^{n/2}}{\Gamma\left(\frac{n}{2}\right)}. \quad (2.57)$$

## 2.2. Geodetikusok

A Riemann-sokaságon a geodetikusok felelnek meg az euklideszi térben megszokott egyeneseknek.

**2.16. Definíció.** Legyen  $(M, \nabla)$  differenciálgeometria. A  $\gamma : \mathbb{R} \rightarrow M$  legalább kétszer differenciálható görbét akkor hívjuk *geodetikusnak*, ha minden  $p \in \text{Ran}(\gamma)$  pont körüli  $\phi$  lokális koordinátázás esetén minden  $1 \leq k \leq \dim M$  indexre

$$\frac{d^2 \gamma^k}{dt^2} + \sum_{i,j=1}^{\dim M} (\Gamma_{ij}^{\cdot k} \circ \gamma) \frac{d\gamma^i}{dt} \frac{d\gamma^j}{dt} = 0 \quad (2.58)$$

teljesül minden  $t \in \text{Dom}(\gamma) \cap \phi^{-1}(U)$  esetén.



**2.10. Tétel.** A  $(\mathcal{P}_1, \nabla)$  differenciálgeometria esetén a  $\gamma$  görbe pontosan akkor geodetikus, ha eleget tesz a

$$\frac{d^2 \gamma(t)}{dt^2} - \frac{(1 - 2\gamma(t))}{2\gamma(t)(1 - \gamma(t))} \left( \frac{d\gamma(t)}{dt} \right)^2 = 0 \quad (2.59)$$

differenciálegyenletnek. Ennek a  $\gamma(0) = a$  és  $\dot{\gamma}(0) = b$  kezdeti feltételhez tartozó megoldása

$$\gamma(t) = \cos^2 \left( -\frac{bt}{2\sqrt{a}\sqrt{1-a}} - \arccos \sqrt{a} \right). \quad (2.60)$$

**2.11. Tétel.** A  $(\mathcal{P}_2, \nabla)$  differenciálgeometria esetén a  $\gamma = (\gamma_1, \gamma_2)$  görbe pontosan akkor geodetikus, ha eleget tesz a

$$\begin{aligned} (1 - \gamma_1(t) - \gamma_2(t)) \frac{d^2 \gamma_1(t)}{dt^2} - \frac{1 - 2\gamma_1(t) - \gamma_2(t) + \gamma_1(t)\gamma_2(t)}{2\gamma_1(t)} \left( \frac{d\gamma_1(t)}{dt} \right)^2 \\ + \gamma_1(t) \frac{d\gamma_1(t)}{dt} \frac{d\gamma_2(t)}{dt} + \frac{\gamma_1(t)(1 - \gamma_1(t))}{2\gamma_2(t)} \left( \frac{d\gamma_2(t)}{dt} \right)^2 = 0 \\ (1 - \gamma_1(t) - \gamma_2(t)) \frac{d^2 \gamma_2(t)}{dt^2} - \frac{1 - 2\gamma_2(t) - \gamma_1(t) + \gamma_1(t)\gamma_2(t)}{2\gamma_2(t)} \left( \frac{d\gamma_2(t)}{dt} \right)^2 \\ + \gamma_2(t) \frac{d\gamma_1(t)}{dt} \frac{d\gamma_2(t)}{dt} + \frac{\gamma_2(t)(1 - \gamma_2(t))}{2\gamma_1(t)} \left( \frac{d\gamma_1(t)}{dt} \right)^2 = 0 \end{aligned} \quad (2.61)$$

differenciálegyenletrendszernek.

**2.12. Tétel.** Legyen  $(p_1, p_2), (q_1, q_2) \in \mathcal{P}_2$  két tetszőleges pont. A két pontot összekötő  $\gamma = (\gamma_1, \gamma_2)$  geodetikus egyenlete

$$\begin{aligned} \gamma_1(t) &= \left( \frac{\sqrt{q_1} - \sqrt{p_1}(\sqrt{p_1q_1} + \sqrt{p_2q_2} + \sqrt{p_3q_3})}{\sqrt{1 - (\sqrt{p_1q_1} + \sqrt{p_2q_2} + \sqrt{p_3q_3})^2}} \sin t + \sqrt{p_1} \cos t \right)^2 \\ \gamma_2(t) &= \left( \frac{\sqrt{q_2} - \sqrt{p_2}(\sqrt{p_1q_1} + \sqrt{p_2q_2} + \sqrt{p_3q_3})}{\sqrt{1 - (\sqrt{p_1q_1} + \sqrt{p_2q_2} + \sqrt{p_3q_3})^2}} \sin t + \sqrt{p_2} \cos t \right)^2. \end{aligned} \quad (2.62)$$

Bevezetve a

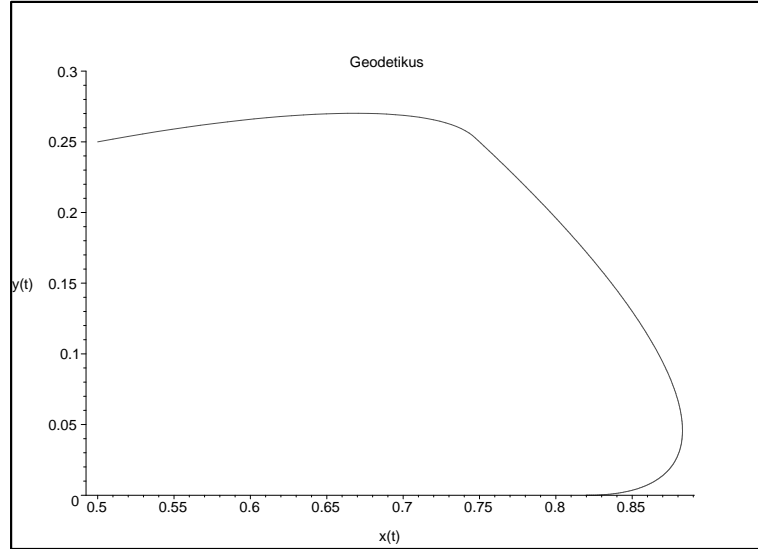
$$t_0 = \arccos(\sqrt{p_1q_1} + \sqrt{p_2q_2} + \sqrt{p_3q_3}) \quad (2.63)$$

paramétert, a geodetikus a

$$\gamma(0) = (p_1, p_2), \quad \gamma(t_0) = (q_1, q_2) \quad (2.64)$$

peremfeltételeknek tesz eleget.

A közönséges Picard–Lindelöf-tétel miatt adott kezdeti feltétel mellett a geodetikusra vonatkozó differenciálegyenletrendszernek van egyértelmű megoldása és a megoldás simán függ a kezdeti feltételtől. A (2.11.) tételben szereplő egyenlet  $\gamma(0) = (1/2, 1/4)$  és  $\dot{\gamma}(0) = (1, 1/5)$  kezdeti feltételhez tartozó megoldását az alábbi ábra szemlélteti, ahol  $x(t) = \gamma_1(t)$  és  $y(t) = \gamma_2(t)$ .



Legyen  $p \in M$ , és legyen  $\phi$  a  $p$  pont  $U$  nyílt környezetének lokális koordinátázása. Az  $(M, g, \nabla)$  sokaságon minden  $v \in T_p M$  vektor esetén létezik olyan  $\varepsilon > 0$  szám, hogy a  $\gamma : ] - \varepsilon, \varepsilon[ \rightarrow M$  geodetikusra  $\gamma(0) = p$ ,  $\dot{\gamma}(0) = v$  és  $\gamma(] - \varepsilon, \varepsilon[) \subseteq U$  teljesül. Legyen  $\mathcal{W}_p \subseteq T_p M$  azon  $v$  érintővektoroknak a halmaza, melyekhez tartozó geodetikusra teljesülnek a fenti feltételek és  $\varepsilon > 1$ .

Ha  $\gamma_1$  a  $v_1$  érintővektorhoz tartozó  $] - \varepsilon_1, \varepsilon_1[$  intervallumon értelmezett geodetikus, valamint  $\gamma_2$  a  $cv_1$  érintővektorhoz tartozó geodetikus, akkor az  $\varepsilon_2 = \frac{\varepsilon_1}{c}$  jelölést használva a  $\gamma_2$  geodetikus egyértelműen kiterjeszthető a  $] - \varepsilon_2, \varepsilon_2[$  intervallumra, és  $\gamma_1(ct) = \gamma_2(t)$  teljesül minden  $t \in ] - \varepsilon_2, \varepsilon_2[$  számra. A geodetikuskok differenciálegyenletrendszerére alkalmazott Picard–Lindelöf-tétel és az iménti megjegyzések alapján a  $\mathcal{W}_p$  halmaz a  $0 \in T_p M$  érintővektor egy környezete, vagyis létezik olyan  $r_p > 0$ , hogy  $S_{r_p} \subseteq \mathcal{W}_p$ , ahol

$$S_{r_p} = \{v \in T_p M \mid g(v, v) < r_p\}. \quad (2.65)$$

**2.17. Definíció.** Legyen  $(M, \nabla)$  differenciálgeometria. A  $p \in M$  ponthoz tartozó *exponenciális leképezést* az

$$\exp_p : S_{r_p} \rightarrow M \quad v \mapsto \gamma_{p,v}(1) \quad (2.66)$$

képlettel értelmezzük, ahol  $\gamma_{p,v}$  a  $\gamma(0) = p$  és  $\dot{\gamma}(0) = v$  feltételeknek eleget tevő geodetikus.

**2.13. Tétel.** Minden  $p \in M$  ponthoz létezik olyan  $\tilde{S}_p \subseteq S_{r_p}$  nyílt halmaz, hogy  $\exp_p$  diffeomorfizmus  $\tilde{S}_p$  és  $\exp(\tilde{S}_p)$  között.

Tekintsük az  $(M, g)$  Riemann-sokaságot a  $\nabla$  Levi–Civita-féle kovariáns deriválással. Minden  $p \in M$  esetén jelölje  $V_n(r)$  az  $\exp(S_r)$  halmaz térfogatát, azaz  $n$ -dimenziós sokaság esetén legyen

$$V_n(r) = \int_{\exp(S_r)} \sqrt{\det(g_{ij})} \quad (0 \leq r < r_p). \quad (2.67)$$

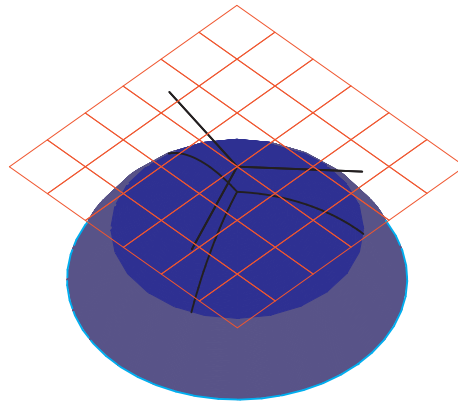
**2.1. Példa.** Legyen az  $M$  sokaság a háromdimenziós euklideszi térben lévő  $a > 0$  sugarú gömbfelszín, melyből (pusztán technikai okokból) elhagyjuk a  $\mathbb{R}_0^- \times \{0\} \times \{0\}$  halmazhoz tartozó elemeket. Az  $M$  sokaság ekkor egyértelműen paraméterezhető a  $(\vartheta, \varphi)$  gömbi koordinátákkal, ahol  $-\pi < \varphi < \pi$  és  $0 < \vartheta < \pi$ . Az euklideszi metrika az érintőtéren ekkor

$$(g_{ij}) = \begin{pmatrix} a^2 & 0 \\ 0 & a^2 \sin^2 \vartheta \end{pmatrix}. \quad (2.68)$$

A  $\gamma(t) = (\vartheta(t), \varphi(t))$  geodetikus egyenlete ekkor

$$\frac{d^2 \vartheta(t)}{dt^2} = \sin \vartheta(t) \cos \vartheta(t) \left( \frac{d\varphi(t)}{dt} \right)^2, \quad \frac{d^2 \varphi(t)}{dt^2} = -2 \frac{\cos \vartheta(t)}{\sin \vartheta(t)} \frac{d\varphi(t)}{dt} \frac{d\vartheta(t)}{dt}. \quad (2.69)$$

A  $\gamma(0) = (\pi/2, 0)$  és  $\dot{\gamma}(0) = (0, v/a)$  kezdeti feltételeknek eleget tevő geodetikus  $\gamma(t) = (\pi/2, tv/a)$ , ahol  $|t| < \frac{\pi a}{|v|}$ . A szimmetria miatt állíthatjuk, hogy a gömbön a geodetikusok a főkörök lesznek. Az exponenciális leképezés ekkor a  $T_{(\pi/2, 0)}M$  érintőtérben lévő  $0 \leq r < \pi a$  sugarú kört a gömb felszínére képezi, létrehozva ott egy  $r$  sugarú gömbre lapított kört. Az exponenciális leképezést az alábbi ábra szemlélteti.



Jelen esetben kétdimenziós sokaságról van szó, ezért  $0 \leq r < \pi a$  esetén

$$V_2(r) = \int_0^{\frac{r}{a}} \int_{-\pi}^{\pi} \sqrt{a^2 \cdot a^2 \sin^2 \vartheta} \, d\varphi \, d\vartheta = 2\pi a^2 \left(1 - \cos \frac{r}{a}\right). \quad (2.70)$$

Kis  $r/a$  érték esetén ez majdnem megegyezik az érintőtérbeli  $r$  sugarú kör területével

$$V_2(r) = r^2\pi - \frac{1}{12} \frac{r^4\pi}{a^2} + \frac{1}{360} \frac{r^6\pi}{a^4} - \frac{1}{20160} \frac{r^8\pi}{a^6} + O\left(\frac{r^{10}}{a^8}\right). \quad (2.71)$$

□

### 2.3. A térfogat Taylor-sora

A  $V_n(r)$  mennyiség „méri”, hogy mekkora a sokaságbeli pont  $r$  sugarú környezete. Első közelítésben gondolhatjuk úgy, hogy ez a térfogat megegyezik az  $n$  dimenziós euklideszi térben lévő  $r$  sugarú gömb térfogatával. Ez az elképzelés – bizonyos értelemben – igaz általános görbült sokaság esetén is. A jelen részben látni fogjuk, hogy a  $V_n(r)$  mennyiség pontosabb közelítésében megjelenik a skalárgörbület. Ez a tény fontos szerepet kap majd a 4.2. részben, ugyanis ezen alapul Petz sejtése.

A  $V_n(r)$  kifejezés sorbafejthető  $r$ -szerint. Először Bertrand, Diguet és Puiseux [13] publikálta 1848-ban a

$$V_n(r) = \pi r^2 \left[ 1 - \frac{K}{12} r^2 + O(r^4) \right] \quad (2.72)$$

kifejezést, mely az  $\mathbb{R}^3$ -ban lévő felületekre érvényes, ahol  $K$  jelöli a Gauss-görbületet. Ezt az eredményt 1917-ben Vermeil [110] és 1939-ben Hotelling [52] általánosította tet-szöleges Riemann-sokaságra. A sorfejtésben lévő következő tagot 1973-ban publikálta Gray [40], majd 1979-ben Gray és Vanhecke újabb tagot írt fel [41]. Az eredményeik bemutatásához először definiáljuk az eddigi mennyiségek átrendezett indexű változatát

$$\begin{aligned} \text{Ric}^{kl} &= \text{Ric}_{ij} g^{ik} g^{jl} & \text{Ric}_i^j &= \text{Ric}_{ik} g^{kj} & R^{ijkl} &= R_{mnop} g^{mi} g^{nj} g^{ok} g^{pl} \\ R_{ijkl}^i &= R_{njkl} g^{ni} & R_{ij}^{..kl} &= R_{ijmn} g^{mk} g^{nl} & R_{..k.l}^{i.j} &= R_{mknl} g^{mi} g^{nj} \end{aligned} \quad (2.73)$$

Ezek segítségével definiáljuk a sorfejtésben szereplő új mennyiségeket

$$\begin{aligned} \|\text{Ric}\|^2 &= \text{Ric}_{ij} \text{Ric}^{ij} & \|R\|^2 &= R_{ijkl} R^{ijkl} & \check{\text{Ric}} &= \text{Ric}_i^j \text{Ric}_j^k \text{Ric}_k^i \\ \langle \text{Ric}, \check{R} \rangle &= \text{Ric}_{ij} R^{ipqr} R_{..pq}^j & \langle \text{Ric} \otimes \text{Ric}, \check{\check{R}} \rangle &= \text{Ric}^{ij} \text{Ric}^{kl} R_{ikjl} & \check{\check{R}} &= R_{ij}^{..kl} R_{..pq}^{kl} R_{pq}^{..ij} \\ \bar{R} &= R_{ikjl} R_{..pq}^{k.l} R^{piqj} \end{aligned} \quad (2.74)$$

Továbbá szükségünk lesz a görbületi mennyiségek kovariáns deriváltjára

$$\begin{aligned}
\|\nabla \text{Scal}\|^2 &= (\partial_i \text{Scal})(\partial_j \text{Scal})g^{ij} & (2.75) \\
(\nabla \text{Ric})_{ijk} &= \partial_k \text{Ric}_{ij} - \Gamma_{ik}^{\cdot p} \text{Ric}_{pj} - \Gamma_{jk}^{\cdot p} \text{Ric}_{ip} \\
\|\nabla \text{Ric}\|^2 &= (\nabla \text{Ric})_{ijk}(\nabla \text{Ric})_{lmn}g^{li}g^{mj}g^{nk} \\
\alpha(\text{Ric}) &= (\nabla \text{Ric})_{jki}(\nabla \text{Ric})_{lmn}g^{li}g^{mj}g^{nk} \\
(\nabla R)_{ijklm} &= \partial_m R_{ijkl} - \Gamma_{im}^{\cdot p} R_{pjkl} - \Gamma_{jm}^{\cdot p} R_{ipkl} - \Gamma_{km}^{\cdot p} R_{ijpl} - \Gamma_{lm}^{\cdot p} R_{ijkm} \\
\|\nabla R\|^2 &= (\nabla R)_{ijklm}(\nabla R)_{nopqr}g^{ni}g^{oj}g^{pk}g^{ql}g^{rm} \\
\Delta \text{Scal} &= (\partial_i \partial_j \text{Scal})g^{ij} - \Gamma_{ij}^{\cdot p} (\partial_p \text{Scal})g^{ij} \\
\langle \Delta \text{Ric}, \text{Ric} \rangle &= (\partial_l (\nabla \text{Ric})_{ijk})g^{kl} \text{Ric}^{ij} - \Gamma_{il}^{\cdot p} (\nabla \text{Ric})_{pjkl}g^{kl} \text{Ric}^{ij} - \\
&\quad - \Gamma_{jl}^{\cdot p} (\nabla \text{Ric})_{ipkl}g^{kl} \text{Ric}^{ij} - \Gamma_{kl}^{\cdot p} (\nabla \text{Ric})_{ijpl}g^{kl} \text{Ric}^{ij} \\
\langle \nabla^2 \text{Scal}, \text{Ric} \rangle &= (\partial_j \partial_i \text{Scal}) \text{Ric}^{ij} - \Gamma_{ij}^{\cdot p} (\partial_p \text{Scal}) \text{Ric}^{ij} \\
\langle \Delta R, R \rangle &= R^{ijkl}g^{mn}(\partial_n (\nabla R)_{ijklm}) - \Gamma_{in}^{\cdot p} (\nabla R)_{pjklm}R^{ijkl}g^{mn} - \\
&\quad - \Gamma_{jn}^{\cdot p} (\nabla R)_{ipklm}R^{ijkl}g^{mn} - \Gamma_{kn}^{\cdot p} (\nabla R)_{ijplm}R^{ijkl}g^{mn} - \\
&\quad - \Gamma_{ln}^{\cdot p} (\nabla R)_{ijkpm}R^{ijkl}g^{mn} - \Gamma_{mn}^{\cdot p} (\nabla R)_{ijklp}R^{ijkl}g^{mn} \\
\Delta^2 \text{Scal} &= (\partial_i \partial_j (\Delta \text{Scal}))g^{ij} - \Gamma_{ij}^{\cdot p} (\partial_p (\Delta \text{Scal}))g^{ij} .
\end{aligned}$$

Tetszőleges  $(M, g)$  Riemann-sokaság esetén a  $V_n(r)$ -re a legtöbb tagot tartalmazó sorfejtést Gray és Vanhecke adta meg [41]. A cikkükben szereplő kifejezés azonban nem adott helyes eredményt ellenőrizhető esetekben. Meglehetősen hosszadalmas számolásukat végigkövetve az alábbi helyes kifejezés adódik a sorfejtésre.

**2.14. Tétel.** Az  $(M, g)$   $n$ -dimenziós Riemann-sokaság  $p$  pontjában a  $V_n(r)$  térfogat  $r$  szerinti Taylor-sora

$$\begin{aligned}
V_n(r) = & \frac{r^n \pi^{n/2}}{\Gamma\left(\frac{n}{2} + 1\right)} \left[ 1 - \frac{Scal}{6(n+2)} r^2 + \frac{-3\|R\|^2 + 8\|Ric\|^2 - 5Scal^2 - 18(\Delta Scal)}{360(n+2)(n+4)} r^4 + \right. \\
& + \frac{1}{720(n+2)(n+4)(n+6)} \left( -\frac{5}{9} Scal^3 - \frac{8}{3} Scal\|Ric\|^2 + Scal\|R\|^2 + \frac{64}{63} \check{Ric} + \right. \\
& + \frac{64}{21} \langle Ric \otimes Ric, \check{R} \rangle + \frac{32}{7} \langle Ric, \dot{R} \rangle + \frac{110}{63} \check{R} + \frac{200}{63} \bar{R} + \frac{45}{7} \|\nabla Scal\|^2 + \frac{45}{14} \|\nabla Ric\|^2 + \\
& + \frac{45}{7} \alpha(Ric) - \frac{45}{14} \|\nabla R\|^2 + 6 Scal(\Delta Scal) + \frac{48}{7} \langle \Delta Ric, Ric \rangle + \frac{54}{7} \langle \nabla^2 Scal, Ric \rangle - \\
& \left. \left. - \frac{30}{7} \langle \Delta R, R \rangle - \frac{45}{7} (\Delta^2 Scal) \right) r^6 + O(r^8) \right]. \tag{2.76}
\end{aligned}$$

A Fisher-metrika esetén a sorfejtés harmadik tagját az alábbiak szerint határozhatjuk meg

$$\begin{aligned}
Ric^{ij} &= \frac{n-1}{4} (-\vartheta_i \vartheta_j + \delta_{ij} \vartheta_i) \tag{2.77} \\
\|Ric\|^2 &= Ric_{ij} Ric^{ij} = n \left( \frac{n-1}{4} \right)^2 \\
\|R\|^2 &= R_{ijkl} R^{ijkl} = \frac{n(n-1)}{8} \\
\Delta Scal &= 0.
\end{aligned}$$

A sorfejtés negyedik tagjában szereplő mennyiségek a következők lesznek

$$\begin{aligned}
\check{Ric} &= n \left( \frac{n-1}{4} \right)^3 & \|\nabla Ric\|^2 &= 0 & \langle \Delta Ric, Ric \rangle &= 0 \tag{2.78} \\
\|\nabla Scal\|^2 &= 0 & \langle \nabla^2 Scal, Ric \rangle &= 0 & \nabla^2 Scal &= 0 \\
\alpha(Ric) &= 0 & \langle Ric \otimes Ric, \check{R} \rangle &= -n \left( \frac{n-1}{4} \right)^3 & \langle Ric, \dot{R} \rangle &= \frac{n}{2} \left( \frac{n-1}{4} \right)^2 \\
\check{R} &= -\frac{n(n-1)}{16} & \bar{R} &= -\frac{n(n-1)(n-2)}{64} & \|\nabla R\|^2 &= 0 \\
\langle \Delta R, R \rangle &= 0.
\end{aligned}$$

**2.15. Tétel.** A  $\mathcal{P}_n$  sokaság esetén a (2.76) sorfejtés a következőképpen specializálódik

$$V_n(r) = \frac{r^n \pi^{n/2}}{\Gamma\left(\frac{n}{2} + 1\right)} \left[ 1 - \frac{n(n-1)}{24(n+2)} r^2 + \frac{n(n-1)(5n-7)}{5760(n+4)} r^4 - \frac{n(n-1)(35n^2 - 112n + 93)}{2903040(n+6)} r^6 + O(r^8) \right]. \quad (2.79)$$

A (2.76) sorfejtés meglehetősen komplikált tagokat is tartalmaz, ezért érdemes ellenőrizni egy jól ismert esetre. Az alábbiakban egy ilyen rövidebb példát mutatunk be.

**2.2. Példa.** Az alapsokaság legyen az (2.1.) példában szereplő  $M$  sokaság. A  $(\vartheta, \phi)$  gömbi koordinátákat használva a metrika

$$g_{11} = a^2, \quad g_{12} = g_{21} = 0, \quad g_{22} = a^2 \sin^2 \theta.$$

A sorfejtésben szereplő segédmennyiségek ekkor könnyen számolhatóak

$$\begin{aligned} \text{Scal} &= \frac{2}{a^2} & \|\text{Ric}\|^2 &= \frac{2}{a^4} & \|R\|^2 &= \frac{4}{a^4} & \Delta \text{Scal} &= 0 & (2.80) \\ \check{\text{Ric}} &= \frac{2}{a^6} & \|\nabla \text{Ric}\|^2 &= 0 & \langle \Delta \text{Ric}, \text{Ric} \rangle &= 0 & \|\nabla \text{Scal}\|^2 &= 0 \\ \langle \nabla^2 \text{Scal}, \text{Ric} \rangle &= 0 & \nabla^2 \text{Scal} &= 0 & \alpha(\text{Ric}) &= 0 & \langle \text{Ric} \otimes \text{Ric}, \check{R} \rangle &= -\frac{2}{a^6} \\ \langle \text{Ric}, \dot{R} \rangle &= \frac{4}{a^6} & \check{R} &= -\frac{8}{a^6} & \bar{R} &= 0 & \|\nabla R\|^2 &= 0 \\ \langle \Delta R, R \rangle &= 0. \end{aligned}$$

A (2.76) alakja ekkor:

$$V(r) = r^2 \pi - \frac{1}{12} \frac{r^4 \pi}{a^2} + \frac{1}{360} \frac{r^6 \pi}{a^4} - \frac{1}{20160} \frac{r^8 \pi}{a^8} + O(r^{10}). \quad (2.81)$$

Ez megegyezik az (2.1.) példában szereplő  $V(r)$  sorfejtéssel.  $\square$

## 2.4. A Fisher-metrika egyértelműsége

Cencov és Morozova mutatták meg, hogy az  $\mathcal{P}_n$  terek közti leképezések segítségével egyértelműen jellemezhető a Fisher-metrika [19, 74]. Ehhez azonban elengedhetetlen a Riemann-sokaságok közti leképezések alaposabb vizsgálata.

**2.18. Definíció.** Legyen  $f : M \rightarrow N$  differenciálható leképezés az  $M$  és  $N$  differenciálható sokaságok között és legyen  $p \in M$ . Adott  $\phi \in \mathcal{F}(N)$  és  $v \in T_p M$  esetén a

$$(f_{*p}(v))(\phi) = v(\phi \circ f) \quad (2.82)$$

leképezés egy  $f_{*p}(v) \in T_{f(p)}N$  vektort definiál. Az  $f$  függvény  $p$  pontbeli *derivált leképezése*

$$f_{*p} : T_p M \rightarrow T_{f(p)}N \quad v \mapsto f_{*p}(v). \quad (2.83)$$

Ha az  $f$  függvény injektív és minden  $p \in M$  pontban a  $f_{*p}$  derivált leképezés is injektív, akkor az  $f$  függvényt *beágyazásnak* nevezzük.

**2.19. Definíció.** Legyen  $M$  és  $N$  differenciálható sokaság. Azt mondjuk, hogy  $N$  *részsokasága*  $M$ -nek, ha  $N \subseteq M$  és az

$$i : N \rightarrow M \quad p \mapsto p \quad (2.84)$$

identikus leképezés minden  $p \in N$  pontbeli  $i_{*p} : T_p N \rightarrow T_p M$  derivált leképezése injektív.

Azt mondjuk, hogy az  $N$  sokaság *1 kodimenziós részsokasága*  $M$ -nek, ha  $N$  pontosan eggyel kisebb dimenziós sokaság, mint  $M$ .

**2.20. Definíció.** Legyen  $f : M \rightarrow N$  differenciálható leképezés az  $M$  differenciálható sokaság és az  $(N, g)$  Riemann-sokaság között. Legyen  $p \in M$ , ekkor a

$$(f^*g)(p) : T_p M \times T_p M \rightarrow \mathbb{R} \quad (x, y) \mapsto g_{f(p)}(f_{*p}(x), f_{*p}(y)) \quad (2.85)$$

leképezés a  $g$ -metrika  $f$ -fel való visszahúzottja.

**2.16. Tétel.** Legyen  $(M, g)$  Riemann-sokaság és  $N$  az  $M$  részsokasága. Ekkor az

$$i : N \rightarrow M \quad p \mapsto p$$

identikus leképezés által visszahúzott  $g$  metrika,  $i^*g$ , Riemann-metrika  $N$ -en.

A fenti tételben szereplő  $i^*g$  metrikát *indukált metrikának* nevezzük.

Legyen  $(M, g)$  Riemann-tér,  $\nabla$  kovariáns deriválás  $M$ -en,  $N \subseteq M$  részsokaság és jelölje  $i : N \rightarrow M$  a beágyazást. Ekkor minden  $p \in N$  esetén  $i_{*p}(T_p N)$  lineáris altere  $T_p M$ -nek. Legyen  $\pi : T_p M \rightarrow T_p N$  ortogonális projekció. Ekkor definiálható a  $\nabla$  kovariáns deriválás visszahúzottja:

$$\nabla_X^{(N)} Y = \pi(\nabla_{i_{*p}(X)} i_{*p}(Y)), \quad (2.86)$$

ahol  $X, Y \in T_p N$  vektorok.

Az alábbi példán megmutatjuk, hogy a  $\mathcal{P}_n$  téren a Fisher-metrika hogyan kapható meg egyszerűen az  $\mathbb{R}^{n+1}$  tér euklideszi metrikájából.



**2.3. Példa.** Tekintsük az

$$f : \mathcal{P}_n \rightarrow \mathbb{R}^{n+1} \quad (\vartheta_1, \dots, \vartheta_n) \mapsto (\sqrt{\vartheta_1}, \dots, \sqrt{\vartheta_n}, \sqrt{1 - \vartheta_1 - \dots - \vartheta_n}) \quad (2.87)$$

leképezést. Legyen  $p \in \mathcal{P}_n$  tetszőleges pont,  $(\partial_{\vartheta_i})_{i=1, \dots, n}$  a  $T_p \mathcal{P}_n$  érintőtér egy bázisa és  $(\partial_i)_{i=1, \dots, n+1}$  a  $T_{f(p)} \mathbb{R}^{n+1}$  tér bázisa. Ekkor  $f$  derivált leképezése a  $p \in \mathcal{P}_n$  pontban

$$f_{*p}(\partial_{\vartheta_i}) = \frac{1}{2\sqrt{\vartheta_i}} \partial_i - \frac{1}{2\sqrt{1 - \vartheta_1 - \dots - \vartheta_n}} \partial_{n+1}. \quad (2.88)$$

Az  $\mathbb{R}^{n+1}$  euklideszi téren a metrikus tenzor  $g_{ik}^{(\mathbb{R}^{n+1})} = \delta_{ik}$ , ennek  $f$ -fel való visszahúzottja

$$(f^*g^{(\mathbb{R}^{n+1})})_{ik} = \frac{1}{4} \left( \delta_{ik} \frac{1}{\vartheta_i} + \frac{1}{1 - \vartheta_1 - \dots - \vartheta_n} \right), \quad (2.89)$$

ami a Fisher-metrika pozitív számszorosa. □

Említettük, hogy a  $\mathcal{P}_n$  tér megfeleltethető az  $n + 1$  dimenziós euklideszi tér egyik hipersíkjának a pozitív ortánsába eső részének. Ha ezt a hipersíkot az origóból induló fénysugarakkal rávetítjük az origó körüli egység sugarú gömbre, akkor az  $n+1$  dimenziós tér euklideszi metrikája éppen a Fisher-metrika számszorosát indukálja a gömbön.

A (1.2.) tétel kimondja, hogy véges halmaz feletti statisztikai modell esetén az alaphalmazon értelmezett leképezés pontosan akkor nem csökkenti a Fisher-mátrixot, ha elégséges statisztikáról van szó. A Fisher-metrika egyértelműségéhez a fenti tulajdonságot kell újra fogalmazni a  $\mathcal{P}$  terek közötti statisztikailag releváns leképezésekre vonatkozóan.

**2.17. Tétel.** *Legyen minden  $n \in \mathbb{N}$  számra  $X_n = \{0, 1, \dots, n\}$ . Tekintsük minden  $n$ -re a  $(\mathcal{P}_n, g_n, \nabla_n)$  hármast, ahol  $g_n$  Riemann-metrika ( $g_n$  tetszőleges Riemann-metrika  $\mathcal{P}_n$ -n) és  $\nabla_n$  kovariáns deriválás (nem feltétlenül Levi-Civita-féle). Jelölje  $g_n^{(F)}$  és  $\nabla_n^{(\alpha)}$  a  $\mathcal{P}_n$  téren a Fisher-metrikát és az  $\alpha$ -kovariáns deriválást. Adott  $f : X_n \rightarrow X_m$  szürjektív leképezésre jelölje  $\tilde{f} : \mathcal{P}_n \rightarrow \mathcal{P}_m$  az  $f$  által indukált leképezést. Adott  $S \subseteq \mathcal{P}_n$  részsokaság esetén legyen  $g^{(S)}$  a  $g_n$ -által indukált metrika  $S$ -en,  $g^{(\tilde{f}(S))}$  a  $g_m$ -által indukált metrika  $\tilde{f}(S)$ -en,  $\nabla^{(S)}$  a  $\nabla_n$  visszahúzottját és  $\nabla^{(\tilde{f}(S))}$  a  $\nabla_m$  visszahúzottját. Ekkor  $\tilde{f}|_S : S \rightarrow \tilde{f}(S)$  differenciálható leképezés. Legyen  $(\tilde{f}|_S)^*g^{(\tilde{f}(S))}$  a  $g^{(\tilde{f}(S))}$  metrika  $\tilde{f}|_S$  leképezéssel való visszahúzottja. Ekkor  $(S, g^{(S)})$ ,  $(\tilde{f}(S), g^{(\tilde{f}(S))})$  és  $(S, (\tilde{f}|_S)^*g^{(\tilde{f}(S))})$  Riemann-tér. Ha minden  $m, n \in \mathbb{N}$  esetén minden  $f : X_n \rightarrow X_m$  szürjektív leképezésre és minden  $S \subseteq \mathcal{P}_n$  olyan részsokaságra, melyre  $\tilde{f} : \mathcal{P}_n \rightarrow \mathcal{P}_m$  elégséges statisztika teljesül, hogy*

$$(\tilde{f}|_S)^*g^{(\tilde{f}(S))} = g^{(S)}, \quad (2.90)$$

akkor létezik pontosan egy olyan  $c > 0$  szám, hogy minden  $n$ -re  $g_n = cg_n^{(F)}$ , továbbá ha minden  $X, Y \in \mathcal{X}(S)$  vektormezőre

$$(\tilde{f}|_S)_* \left( \nabla_X^{(S)} Y \right) = \nabla_{(\tilde{f}|_S)_*(X)}^{(\tilde{f}|_S)^*} (\tilde{f}|_S)_*(Y) \quad (2.91)$$

teljesül, akkor létezik olyan  $\beta \in [-1, 1]$  és  $d > 0$  paraméter, hogy minden  $n$ -re  $\nabla_n = d\nabla_n^{(\beta)}$  teljesül.

A tétel szerint a Fisher-metrika és az  $\alpha$  kovariáns deriválások egyértelműek véges halmazokon pozitív számszorzó erejéig.

## 2.5. Részsokaság skalárgörbülete

A 4.1. részben meghatározzuk a kvantummechanikai állapotter skalárgörbületét. Ott úgy tekintjük majd az állapotteret, mint egy nagyobb sokaság 1-kodimenziós részsokasága. Ezért fontos a nagyobb sokaság skalárgörbülete és a részsokaság skalárgörbülete közötti kapcsolatot tisztázni. Továbbá a jelen részben bemutatjuk az állapotteren történő számolás klasszikus megfelelőjét.

Megmutatjuk, hogy a  $\mathcal{P}_n$  tér természetes módon megkapható, mint egy lapos tér részsokasága. Ehhez definiáljuk a

$$\tilde{\mathcal{P}}_n := \{p(x, \vartheta_0, \dots, \vartheta_n) \mid \forall i \in \{0, \dots, n\} : \vartheta_i \in \mathbb{R}\} \quad (2.92)$$

teret. A  $\tilde{\mathcal{P}}_n$  tér differenciálható sokaság, és minden  $p \in \tilde{\mathcal{P}}$  pont esetén  $(\partial_{\vartheta_i})_{i=0, \dots, n}$  a  $T_p \tilde{\mathcal{P}}$  érintőtér bázisa. A

$$\tilde{g}_{ik} = \delta_{ik} \frac{1}{\vartheta_i} \quad (2.93)$$

metrikus tenzor Riemann-metrikát definiál a  $\tilde{\mathcal{P}}_n$  téren.

**2.18. Tétel.** A  $(\tilde{\mathcal{P}}_n, \tilde{g})$  Riemann-sokaságon a Levi-Civita-féle kovariáns deriválás a

$$\tilde{\Gamma}_{ij}^{\cdot k} = -\frac{1}{2} \frac{1}{\vartheta_i} \delta_{ij} \delta_{jk} \quad (2.94)$$

másodfajú Christoffel-szimbólummal jellemezhető.

A továbbiakban  $\tilde{\nabla}$  jelöli ezt a kovariáns deriválást.

**2.19. Tétel.** A  $(\tilde{\mathcal{P}}_n, \tilde{g}, \tilde{\nabla})$  tér Riemann-féle görbületi tenzora

$$\tilde{R}_{ijk}^{\cdot l} = 0. \quad (2.95)$$

Tehát a  $(\tilde{\mathcal{P}}_n, \tilde{g}, \tilde{\nabla})$  tér lapos.

**2.20. Tétel.** A  $\gamma : \mathbb{R} \rightarrow \tilde{\mathcal{P}}_n$  görbe akkor geodetikusa a  $(\tilde{\mathcal{P}}_n, \tilde{g}, \tilde{\nabla})$  térnek, ha eleget tesz a

$$\forall k \in \{1, \dots, n\} : \quad \frac{d^2 \gamma_k(t)}{dt^2} - \frac{1}{2} \left( \frac{d\gamma_k(t)}{dt} \right)^2 = 0 \quad (2.96)$$

differenciálegyenletrendszernek. Ennek a  $\gamma(0) = a$  és a  $\dot{\gamma}(0) = b$  kezdeti feltételnek eleget tevő megoldása

$$\forall k \in \{1, \dots, n\} : \quad \gamma_k(t) = \frac{b_k^2}{4a_k} t^2 + b_k t + a_k . \quad (2.97)$$

**2.21. Tétel.** Tekintsük az  $(\mathcal{P}_n, g_n)$  és  $(\tilde{\mathcal{P}}_n, \tilde{g}_n)$  Riemann-sokaságokat. A

$$f : \mathcal{P}_n \rightarrow \tilde{\mathcal{P}}_n \quad p(x, \vartheta_0, \dots, \vartheta_n) \mapsto p(x, \vartheta_0, \dots, \vartheta_n) \quad (2.98)$$

beágyazással  $\mathcal{P}_n$  a  $\tilde{\mathcal{P}}_n$  részsokasága. A  $\tilde{g}_n$  metrika  $f$ -fel való visszahúzottja megegyezik a  $g_n$  metrikával.

A továbbiakban megmutatjuk, hogyan számolható ki  $\mathcal{P}_n$  skalárgörbülete a  $\tilde{\mathcal{P}}_n$  tér segítségével.

**2.21. Definíció.** Legyen  $(M, g)$  Riemann-sokaság és  $N \subset M$  az  $M$  sokaság 1-kodimenziós részsokasága az  $f : N \rightarrow M$  beágyazással. Azt mondjuk, hogy a

$$n : N \rightarrow TM \quad p \mapsto n(p) \quad (2.99)$$

leképezés az  $N$  sokaság normálvektormezője (vagy  $N$  felületi merőlegese), ha minden  $p \in N$  esetén  $n(p) \in T_{f(p)}M$ ,  $g_p(n(p), n(p)) = 1$  és  $n(p) \perp f_{*p}(T_pN)$  teljesül.

A Riemann-tér minden 1-kodimenziós részsokaságának létezik normálvektormezője. A továbbiakban olyan részsokaságokról lesz szó, ahol a normálvektormező-leképezés kellően sokszor differenciálható.

**2.22. Tétel.** Legyen  $(\tilde{M}, \tilde{g})$   $n+1$  dimenziós Riemann-sokaság,  $M \subset \tilde{M}$  1-kodimenziós részsokasága  $M$ -nek,  $g$  a  $\tilde{g}$  által indukált metrika  $M$ -en,  $\tilde{\nabla}$  a  $\tilde{g}$  metrikához tartozó Levi-Civita-féle kovariáns deriválás és  $n : M \rightarrow T\tilde{M}$  az  $M$  sokaság normálvektormezője. Minden  $X, Y \in TM$  esetén definiáljuk az

$$S(X, Y) : M \rightarrow \mathbb{R} \quad p \mapsto -\tilde{g}_p(\tilde{\nabla}_X n, Y) \quad (2.100)$$

függvényt. Ekkor  $X, Y, Z, U \in \mathcal{X}(M)$  esetén

$$\tilde{g}(\tilde{R}(X, Y)Z, U) = g(R(X, Y)Z, U) + S(X, Y)S(Z, U) - S(Z, Y)S(X, U) \quad (2.101)$$

teljesül, ahol  $R$  és  $\tilde{R}$  jelöli az  $(M, g)$  illetve az  $(\tilde{M}, \tilde{g})$  terek görbületesi tenzorát.

Továbbá, ha  $(A_t)_{t=1,\dots,n}$  a  $T_p M$  érintőtér ortonormált bázisa (azaz  $g(A_t, A_s) = \delta_{ts}$ ), akkor az  $(M, g)$  tér  $p$  pontbeli  $\text{Scal}(p)$  skalárgörbületére

$$\text{Scal}(p) = \sum_{t,s=1}^n \tilde{g}(\tilde{R}(A_t, A_s)A_s, A_t) + S(A_s, A_s)S(A_t, A_t) - S(A_t, A_s)S(A_s, A_t) \quad (2.102)$$

teljesül.

Adott  $X, Y \in T_p M$  vektorokra is értelmezhető az  $S(X, Y)$  mennyiség, mert a 2.2. tétel alapján lehet értelmezni a vektor szerinti kovariáns deriválást.

A  $(\mathcal{P}_n, g)$  Riemann-tér skalárgörbületét is egyszerűen meg lehet határozni a (2.102) képlet segítségével.

**2.4. Példa.** A  $(\mathcal{P}_n, g)$  tér skalárgörbülete: A (2.21.) tétel alapján  $\mathcal{P}_n$  1 kodimenziós részsokasága a  $(\tilde{\mathcal{P}}_n)$  sokaságnak. A  $\mathcal{P}_n$  normálvektormezőjét adja a

$$n : \mathcal{P}_n \rightarrow T\tilde{\mathcal{P}}_n \quad (\vartheta_0, \dots, \vartheta_n) \mapsto \vartheta_0 \partial_0 + \dots + \vartheta_n \partial_n \quad (2.103)$$

leképezés, ugyanis minden  $i = 1, \dots, n$  esetén a  $\partial_i - \partial_0 \in T\mathcal{P}_n$  vektormező esetén

$$\tilde{g}(\partial_i - \partial_0, n) = \tilde{g}(\partial_i - \partial_0, \vartheta_0 \partial_0 + \dots + \vartheta_n \partial_n) = 0 \quad (2.104)$$

teljesül.

A (2.98) képlettel értelmezett  $f$  beágyazás érintőleképezésére bevezetjük a  $\tilde{\partial} = f_*(\partial)$  jelölést. A (2.100) képlettel értelmezett  $S$  leképezés a  $\partial_i, \partial_j \in T\mathcal{P}_n$  ( $i, j = 1, \dots, n$ ) vektorokon az

$$\begin{aligned} S(\partial_i, \partial_j) &= -\tilde{g}(\tilde{\nabla}_{-\tilde{\partial}_0 + \tilde{\partial}_i} n, -\tilde{\partial}_0 + \tilde{\partial}_j) = -\tilde{g}\left(\frac{-1}{2}\tilde{\partial}_0 + \frac{1}{2}\tilde{\partial}_i, -\tilde{\partial}_0 + \tilde{\partial}_j\right) \\ &= \frac{-1}{2\vartheta_i}\delta_{ij} - \frac{1}{2\vartheta_0} = \frac{-1}{2}g(\partial_i, \partial_j) \end{aligned} \quad (2.105)$$

értéket veszi fel.

Legyen  $\vartheta \in \mathcal{P}_n$  pont esetén  $(A_i)_{i=1,\dots,n}$  ortonormált rendszer a  $T_\vartheta \mathcal{P}_n$  érintőtérben. A (2.19.) tétel alapján a  $(\tilde{\mathcal{P}}_n, \tilde{g})$  tér lapos, ezért a skalárgörbületet kifejező (2.102) tétel ebben az esetben így alakul

$$\text{Scal} = \sum_{t,s=1}^n S(A_s, A_s)S(A_t, A_t) - S(A_t, A_s)S(A_s, A_t) = \sum_{t,s=1}^n \frac{1}{4} - \delta_{ts} \frac{1}{4} = \frac{n(n-1)}{4}. \quad (2.106)$$

Ez pedig megegyezik a (2.50) képlettel. □

## 2.6. A normális eloszlások geometriája

Ebben a fejezetben a többváltozós normális eloszlás geometriájával foglalkozunk részletesen. Ebből meglepően sok részletet tudunk majd jól hasznosítani a nemkommutatív statisztikai vizsgálatok során. Később (a 4.1. részben) meghatározzuk a kvantummechanikai állapottér skalárgömbületét és látni fogjuk, hogy a normális eloszlások geometriája és az állapottér geometriája bizonyos értelemben hasonló szerkezetű. Azért, hogy ez a hasonlóság feltűnőbb legyen egy ponton kicsit bonyolultabb formában írunk fel egy összefüggést, és a jelen részen keresztül ezzel a bonyolultabb formulával számolunk. Ennek az előnye az, hogy a kvantummechanikai állapottér skalárgömbületének a meghatározásánál ezeket a számolásokat már nem kell elvégezni, csak kiegészíteni újabbakkal.

Legyen  $n$  természetes szám, az alaphalmaz  $X = \mathbb{R}^n$ , és a paramétertér a valós  $n \times n$ -es szimmetrikus pozitív definit mátrixok halmaza, vagyis

$$M_n^+ = \{D \in M_n(\mathbb{R}) \mid D = D^*, D > 0\} . \quad (2.107)$$

Az  $S$  halmaz elemei legyenek az

$$f : M_n^+ \times X \rightarrow \mathbb{R} \quad (D, \underline{x}) \mapsto f(D, \underline{x}) = \frac{\sqrt{\det D}}{\sqrt{(2\pi)^n}} \exp\left(-\frac{1}{2}\langle \underline{x}, D\underline{x} \rangle\right) \quad (2.108)$$

paraméteres sűrűségfüggvények. A továbbiakban az  $(X, S, M_n^+)$  statisztikai modell Fisher-féle információs mátrixa által indukált Riemann-geometriáját elemezzük. A jelen fejezetben az  $(X, S, M_n^+)$   $n$ -dimenziós normális eloszlás statisztikai modellje kifejezésen az imént bevezetett statisztikai modellt értjük.

Az  $M_n^+$  paramétertér az önadjungált mátrixok halmazának nyílt részhalmaza, vagyis az  $\mathbb{R}^{n(n+1)/2}$  tér nyílt részhalmaza. Így természetes módon (egy térképpel lefedhető) differenciálható sokaság. Minden  $D \in M_n^+$  pont esetén az érintőtér azonosítható az  $n \times n$ -es, valós, önadjungált mátrixok halmazával, vagyis

$$T_D M_n^+ = \{X \in M(\mathbb{R}, n) \mid X = X^*\} . \quad (2.109)$$

Tetszőleges  $X \in T_D M_n^+$  vektor esetén létezik olyan  $\varepsilon > 0$ , hogy a

$$\gamma : ]-\varepsilon, \varepsilon[ \rightarrow M(\mathbb{R}, n) \quad t \mapsto D + tX \quad (2.110)$$

leképezésre  $\text{Ran } \gamma \subseteq M_n^+$  teljesül. Az  $X \in T_D M_n^+$  vektor az  $f \in \mathcal{F}(M_n^+)$  függvényhez az

$$X(f) = \left. \frac{df(D + tX)}{dt} \right|_{t=0} \quad (2.111)$$

értéket rendeli.

Legyen  $D \in M_n^+$  és  $X \in T_D M_n^+$ . Ekkor az  $f(D, \underline{x}) \in S$  függvény  $X$ -irányú deriváltját a

$$\frac{\partial f(D, \underline{x})}{\partial X} = \left. \frac{\partial f(D + tX, \underline{x})}{\partial t} \right|_{t=0} \quad (2.112)$$

képlettel értelmezzük.

Az  $M_n^+$  paraméterter, és bármely pontbeli érintőtér is,  $\frac{n(n+1)}{2}$  dimenziós. A könnyebb számolás kedvéért  $1 \leq i, j \leq n$  indexek esetén definiáljuk az

$$(E_{ij})_{ab} = \delta_{ia} \delta_{jb} \quad (1 \leq a, b \leq n) \quad (2.113)$$

mátrixegységet, valamint  $1 \leq i \leq j \leq n$  esetén az

$$F_{ij} = E_{ij} + E_{ji} \quad (2.114)$$

mátrixokat. Ekkor az  $(F_{ij})_{1 \leq i < j \leq n}$  vektorhalmaz bármely pont feletti érintőtér bázisa.

Adott  $D \in M_n^+$  pontbeli differenciálgeometriai vizsgálódásoknál az általánosság megszorítása nélkül feltehetjük, hogy a  $D$  mátrix diagonális, főátlójában a sajátértékei szerepelnek. (Alkalmas unitér transzformációval minden önadjungált mátrix diagonalizálható és az unitér transzformáció által meghatározott új koordinátarendszerben dolgozunk tovább.)

A Fisher-féle információs mátrix meghatározását az alábbi tétel segíti.

**2.23. Tétel.** *Legyen az  $(X, S, M_n^+)$  hármas  $n$ -dimenziós normális eloszlás statisztikai modellje,  $D \in M_n^+$  tetszőleges pont és  $X, Y \in T_D M_n^+$  érintőtérbeli vektor. Ekkor a  $D \in M_n^+$  pontban a Fisher-féle információs mátrix által indukált Riemann-metrikára*

$$g^{(F)}(D)(X, Y) = \int_{\mathbb{R}^n} \frac{1}{f(D, \underline{x})} \frac{\partial f(D, \underline{x})}{\partial X} \frac{\partial f(D, \underline{x})}{\partial Y} d\underline{x} = \frac{1}{2} \text{Tr}(D^{-1} X D^{-1} Y) \quad (2.115)$$

teljesül.

*Bizonyítás.* A tétel előtti megjegyzés értelmében legyen  $D = \sum_{k=1}^n \lambda_k E_{kk}$  alakú és legyen tetszőleges  $1 \leq i, j \leq n$  és  $1 \leq k, l \leq n$  indexre  $E_{ij}$  és  $E_{kl}$  a mátrixegység. Ekkor

$$g^{(F)}(D)(E_{ij}, E_{kl}) = \int_{\mathbb{R}^n} f(D, \underline{x}) \frac{\partial \ln f(D, \underline{x})}{\partial E_{ij}} \frac{\partial \ln f(D, \underline{x})}{\partial E_{kl}} d\underline{x} \quad (2.116)$$

teljesül. A parciális deriváltakat az alábbiaknak megfelelően alakítjuk át.

$$\frac{\partial \ln f(D, \underline{x})}{\partial E_{ij}} = \frac{1}{2} \frac{\partial \ln \det D}{\partial E_{ij}} - \frac{1}{2} \frac{\partial \langle \underline{x}, D \underline{x} \rangle}{\partial E_{ij}} = \frac{1}{2} (D^{-1})_{ji} - \frac{1}{2} x_i x_j \quad (2.117)$$

Ezek alapján

$$g^{(F)}(D)(E_{ij}, E_{kl}) = \frac{1}{4}\delta_{ij}\delta_{kl}\frac{1}{\lambda_i\lambda_k} - \frac{1}{4}\delta_{ij}\frac{1}{\lambda_i}\int_{\mathbb{R}^n} f(D, \underline{x})x_kx_l \, d\underline{x} - \frac{1}{4}\delta_{kl}\frac{1}{\lambda_k}\int_{\mathbb{R}^n} f(D, \underline{x})x_ix_j \, d\underline{x} + \frac{1}{4}\int_{\mathbb{R}^n} f(D, \underline{x})x_ix_jx_kx_l \, d\underline{x}. \quad (2.118)$$

Parciális integrálásokkal a

$$g^{(F)}(D)(E_{ij}, E_{kl}) = \frac{1}{4}\delta_{ij}\delta_{kl}\frac{1}{\lambda_i\lambda_k}(3\delta_{ik} - 1) \quad (2.119)$$

kifejezést adódik. Az  $F_{ij}$  és  $F_{kl}$  vektorokra, az összetett függvény deriválási szabálya alapján a

$$\begin{aligned} g^{(F)}(D)(F_{ij}, F_{kl}) &= g^{(F)}(D)(E_{ij}, E_{kl}) + g^{(F)}(D)(E_{ij}, E_{lk}) + \\ &+ g^{(F)}(D)(E_{ji}, E_{kl}) + g^{(F)}(D)(E_{ji}, E_{lk}) \\ &= \frac{1}{\lambda_i\lambda_j}(\delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk}) \end{aligned} \quad (2.120)$$

kifejezést kapjuk. Az

$$\frac{1}{\lambda_i\lambda_j}(\delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk}) = \frac{1}{2}\text{Tr} D^{-1}F_{ij}D^{-1}F_{kl} \quad (2.121)$$

azonosság pedig könnyen ellenőrizhető. A (2.115) egyenlet teljesül az érintőtérbeli  $(F_{ij})_{1 \leq i \leq j \leq n}$  báziselemekre, a leképezések linearitása miatt ekkor a (2.115) egyenlet minden  $X, Y \in T_D M_n^+$  érintővektorra is teljesül.  $\square$

Ennek a tételnek a segítségével egyszerűen igazolható, hogy a  $g^{(F)}$  leképezés minden  $D \in M_n^+$  pont esetén pozitív definit. Tehát  $g^{(F)}$  Riemann-metrikát határoz meg az  $M_n^+$  sokaságon.

Mivel csak a  $g^{(F)}$  Fisher-féle metrikával foglalkozunk ebben a fejezetben, a  $g = 2g^{(F)}$  jelölést használjuk a továbbiakban. Az  $(M_n^+, g)$  tér Riemann-geometriájának néhány fontosabb differenciálgeometriai mennyiségét határozzuk meg.

A továbbiakban azt az egyszerűsítő jelölést követjük, hogy amennyiben  $A : V \rightarrow V$  lineáris leképezés és  $c \in \mathbb{C}$  paraméter, akkor az  $A + c \text{id}_V$  leképezést röviden  $(A + c)$ -nek írjuk.

A (2.115) képletre az alábbi ekvivalens alakot is fogjuk használni

$$\text{Tr}(D^{-1}XD^{-1}Y) = \text{Tr} \int_T (D + t)^{-1}X(D + t)^{-1}Y \, d\mu(t), \quad (2.122)$$

ahol  $T = [0, \infty[$  és  $\mu = \delta_0$ .

Ezzel a lépéssel elértük, hogy a metrika kifejezhető egy speciális integrálformulával. (Az állapottéren értelmezett metrikák is ilyen alakúak lesznek.)

A  $T_D M_n^+$  érintőtéren vezessük be a *Hilbert–Schmidt-féle skaláris szorzást* (vagy *kanonikus skaláris szorzást*)

$$\langle \cdot, \cdot \rangle : T_D M_n^+ \times T_D M_n^+ \rightarrow \mathbb{R} \quad (X, Y) \mapsto \langle X, Y \rangle = \text{Tr}(XY), \quad (2.123)$$

valamint a

$$G : M_n^+ \rightarrow \text{Lin}(T_D M_n^+, T_D M_n^+) \quad D \mapsto (X \mapsto G(D)(X)) \quad (2.124)$$

$$G(D)(X) = \int_T (D+t)^{-1} X (D+t)^{-1} d\mu(t)$$

leképezést. Ekkor a Fisher-metrikára

$$g(D)(X, Y) = \langle G(D)(X), Y \rangle = \langle X, G(D)(Y) \rangle \quad (2.125)$$

teljesül.

Legyen  $D = \sum_{k=1}^n \lambda_k E_{kk}$  alakú. Az  $F_{ij}$  ( $1 \leq i \leq j \leq n$ ) vektorok  $G(D)$  leképezés általi képére minden  $1 \leq a, b \leq n$  esetén

$$\begin{aligned} (G(D)(F_{ij}))_{ab} &= \int \sum_{k,l=1}^n ((t+D)^{-1})_{ak} (F_{ij})_{kl} ((t+D)^{-1})_{lb} d\mu(t) = \\ &= \int_T (t+\lambda_a)^{-1} (F_{ij})_{ab} (t+\lambda_b)^{-1} d\mu(t) \end{aligned} \quad (2.126)$$

teljesül. Definiáljuk az

$$m : \mathbb{R}^+ \times \mathbb{R}^+ \rightarrow \mathbb{R} \quad (x, y) \mapsto \int_T (x+t)^{-1} (y+t)^{-1} d\mu(t) \quad (2.127)$$

függvényt. Ha a diagonális  $D$  mátrix a  $\sum_{k=1}^n \lambda_k E_{kk}$  alakba írható, akkor  $1 \leq i, j \leq n$  index esetén az

$$m_{ij} = m(\lambda_i, \lambda_j) \quad (2.128)$$

rövidítést használjuk. A jelen esetben

$$m_{ij} = \frac{1}{\lambda_i \lambda_j}. \quad (2.129)$$

Ezzel a jelöléssel a (2.126) egyenlet az

$$G(D)(F_{ij}) = m_{ij} F_{ij} \quad (2.130)$$

alakra egyszerűsödik.

Az  $m$  függvénnyel könnyen kifejezhető a Fisher-féle metrika.



**2.24. Tétel.** Legyen a  $D \in M_n^+$  diagonális mátrix  $D = \sum_{k=1}^n \lambda_k E_{kk}$  alakú. Ekkor minden  $1 \leq i \leq j \leq n$  és  $1 \leq k \leq l \leq n$  indexpárra

$$g(F_{ij}, F_{kl}) = \begin{cases} 0 & \text{ha } (i, j) \neq (k, l) \\ 2m_{ij} & \text{ha } i = k < j = l \\ 4m_{ii} & \text{ha } i = k = j = l \end{cases} \quad (2.131)$$

teljesül.

A további számolások előtt említsük meg, hogy  $E, F$  Banach-terek esetén az  $f : E \rightarrow F$  kétszer differenciálható függvény deriváltjait a

$$d^0 f : E \rightarrow F \quad a \mapsto f(a) \quad (2.132)$$

$$df : E \rightarrow \text{Lin}(E, F) \quad a \mapsto (x \mapsto df(a)(x))$$

$$d^2 f : E \rightarrow \text{Lin}(E, \text{Lin}(E, F)) \simeq \text{Lin}(E^2, F) \quad a \mapsto (x \mapsto (y \mapsto d^2 f(a)(x)(y)))$$

alakban írjuk fel.

Ennek megfelelően a

$$g : M_n^+ \rightarrow \text{Lin}((TM_n^+)^2, \mathbb{R}) \quad D \mapsto ((X, Y) \mapsto g(D)(X, Y)) \quad (2.133)$$

Riemann-metrika deriváltja

$$dg : M_n^+ \rightarrow \text{Lin}(TM_n^+ \rightarrow \text{Lin}((TM_n^+)^2, \mathbb{R})) \quad (2.134)$$

$$D \mapsto (Z \mapsto ((X, Y) \mapsto dG(D)(Z)(X, Y)))$$

leképezés, a második deriváltja pedig

$$d^2 g : M_n^+ \rightarrow \text{Lin}(TM_n^+ \mapsto \text{Lin}(TM_n^+ \rightarrow \text{Lin}((TM_n^+)^2, \mathbb{R}))) \quad (2.135)$$

$$D \mapsto (V \mapsto (Z \mapsto ((X, Y) \mapsto dG(D)(V)(Z)(X, Y)))) .$$

Adott  $X, Y$  vektormezők esetén a kovariáns deriválásra a  $\Gamma(X, Y) = \nabla_X Y$  konvenciót alkalmazzuk. A (2.3.) tételnek megfelelően egyértelműen létezik egy  $\Gamma$  Levi-Civita-féle kovariáns deriválás az  $M_n^+$  sokaságon, melyet a (2.31) képlet alapján a  $g$  Fisher-metrikával a

$$g(D)(\Gamma(D)(X, Y), Z) = \frac{1}{2} \left( dg(D)(X)(Y, Z) + dg(D)(Y)(X, Z) - dg(D)(Z)(X, Y) \right) \quad (2.136)$$

formában fejezhetünk ki.

A  $dg$  leképezés meghatározásához előbb határozzuk meg az

$$i : M_n \rightarrow M_n \quad a \mapsto i(a) = a^{-1} \quad (2.137)$$

inverzképezés deriváltját valamely  $a \in \mathcal{M}_n^+$  pontban. Ha  $b \in M_n$  operátorra

$$\|b\| < \frac{1}{\|a^{-1}\|} \quad (2.138)$$

teljesül, akkor az  $a + b$  operátor is invertálható. Ugyanis ekkor  $\|a^{-1}(-b)\| < 1$  teljesül, és az operátorok inverzére vonatkozó Neumann-tétel értelmében az  $c = 1 - a^{-1}(-b)$  operátor invertálható. Mivel  $c = a^{-1}(a + b)$ , ezért  $a + b$  is invertálható.

Az

$$i(a + b) - i(a) = -(a + b)^{-1}ba^{-1} \quad (2.139)$$

egyenlőség miatt az  $i(a + b) - i(a)$  mennyiségben a  $b$  változó szerinti nullad- és elsőrendű tagok az alábbiak.

$$-a^{-1}ba^{-1} \quad (2.140)$$

Ezek alapján

$$di : M_n \rightarrow \text{Lin}(M_n, M_n) \quad a \mapsto (x \mapsto di(a)(x)) , \quad (2.141)$$

ahol

$$di(a)(x) = -a^{-1}xa^{-1} . \quad (2.142)$$

Az inverzképezés deriváltja és a szorzatfüggvény deriválási szabálya alapján, a  $G$  leképezés deriváltja

$$dG : M_n^+ \rightarrow \text{Lin}(TM_n^+, \text{Lin}(TM_n^+, TM_n^+)) \quad D \mapsto \left( Y \mapsto (X \mapsto dG(D)(Y)(X)) \right) \quad (2.143)$$

$$dG(D)(Y)(X) = - \int_T (D + t)^{-1}X(D + t)^{-1}Y(D + t)^{-1} + (D + t)^{-1}Y(D + t)^{-1}X(D + t)^{-1} d\mu(t) .$$

Legyen  $D = \sum_{k=1}^n \lambda_k E_{kk}$  alakú. Az  $E_{ij}$  ( $1 \leq i, j \leq n$ ) és  $E_{kl}$  ( $1 \leq k, l \leq n$ ) mátrixegységeken a  $G(D)$  leképezés deriváltja  $1 \leq a, b \leq n$  esetén

$$\left( dG(D)(E_{ij})(E_{kl}) \right)_{ab} = -(\delta_{ia}\delta_{jk}\delta_{lb} + \delta_{ak}\delta_{li}\delta_{jb}) \int_T (t + \lambda_i)^{-1}(t + \lambda_j)^{-1}(t + \lambda_k)^{-1} d\mu(t). \quad (2.144)$$

Definiáljuk az

$$m : \mathbb{R}^+ \times \mathbb{R}^+ \times \mathbb{R}^+ \rightarrow \mathbb{R} \quad (x, y, z) \mapsto \int_T (x+t)^{-1}(y+t)^{-1}(z+t)^{-1} d\mu(t) \quad (2.145)$$

függvényt. Ha a diagonális  $D$  mátrix a  $D = \sum_{k=1}^n \lambda_k E_{kk}$  alakba írható, akkor  $1 \leq i, j \leq n$  index esetén legyen

$$m_{ijk} = m(\lambda_i, \lambda_j, \lambda_k). \quad (2.146)$$

A jelen esetben

$$m_{ijk} = \frac{1}{\lambda_i \lambda_j \lambda_k}. \quad (2.147)$$

Ezekkel a jelölésekkel könnyen kifejezhető a  $G$  leképezés deriváltja; a (2.144) egyenlet új alakja

$$dG(D)(E_{ij})(E_{kl}) = -(m_{ijl}\delta_{jk}E_{il} + m_{jkl}\delta_{il}E_{kj}) \quad (2.148)$$

lesz. A  $dG(D)$  leképezés az  $(F_{ij})_{1 \leq i < j \leq n}$  báziselemeket az alábbi módon transzformálja.

$$dG(D)(F_{ij})(F_{kl}) = -(F_{il}m_{ijl}\delta_{jk} + F_{jk}m_{ijk}\delta_{il} + F_{ik}m_{ijk}\delta_{jl} + F_{jl}m_{ijl}\delta_{ik}) \quad (2.149)$$

A  $g$  Fisher-metrika deriváltja

$$dg : M_n^+ \rightarrow \text{Lin}\left(TM_n^+, \text{Lin}(TM_n^+ \times TM_n^+, \mathbb{R})\right) \quad (2.150)$$

$$D \mapsto \left( Z \mapsto ((X, Y) \mapsto dg(D)(Z)(X, Y)) \right)$$

$$dg(D)(Z)(X, Y) = \langle dg(D)(Z)(X), Y \rangle = -\text{Tr} \int_T (D+t)^{-1} Z (D+t)^{-1} X (D+t)^{-1} Y + \\ + (D+t)^{-1} X (D+t)^{-1} Z (D+t)^{-1} Y d\mu(t).$$

A Levi–Civita-féle kovariáns deriválásra a (2.136) egyenletből

$$g(D)\left(\Gamma(D)(X, Y), Z\right) = \quad (2.151) \\ = -\frac{1}{2} \text{Tr} \int_T (D+t)^{-1} (X(D+t)^{-1} Y + Y(D+t)^{-1} X) (D+t)^{-1} Z d\mu(t)$$

adódik. Ezek alapján a  $\Gamma(D)(X, Y)$  vektor  $G(D)$  transzformáltjára az (2.125) egyenlet alapján

$$g(D)\left(\Gamma(D)(X, Y), Z\right) = \left\langle G(D)\left(\Gamma(D)(X, Y)\right), Z \right\rangle \quad (2.152)$$

teljesül. Ebből

$$G(D)(\Gamma(D)(X, Y)) = \frac{1}{2}dG(D)(X)(Y) \quad (2.153)$$

adódik.

A  $\Gamma$  Levi–Civita-féle kovariáns deriválás kiszámításához ismerni kell a  $G(D)$  leképezés inverzét. Ehhez vezessük be a

$$G^{(-1)} : M_n^+ \rightarrow \text{Lin}(TM_n^+, TM_n^+) \quad D \mapsto \left( X \mapsto G^{(-1)}(D)(X) = (G(D))^{-1}(X) \right) \quad (2.154)$$

függvényt. Legyen  $D = \sum_{k=1}^n \lambda_k E_{kk}$  alakú és  $F_{ij}$  ( $1 \leq i \leq j \leq n$ ) vektor. A (2.130) egyenletből

$$G^{(-1)}(D)(F_{ij}) = G(D)^{-1}(F_{ij}) = \frac{1}{m_{ij}} F_{ij} \quad (2.155)$$

adódik. Tehát a  $G(D)$  leképezés és az inverze is csak számmal szorozza az  $F_{ij}$  bázisvektorokat. A jelen esetben általános  $X \in T_D M$  vektorra is felírhatjuk a  $G(D)$  leképezést és inverzét

$$G(D)(X) = D^{-1} X D^{-1} \quad G^{(-1)}(D)(X) = D X D. \quad (2.156)$$

Ezek alapján a Levi–Civita-féle kovariáns deriválásra adott (2.153) kifejezés így alakítható.

$$\Gamma : M_n^+ \rightarrow \text{Lin}(TM_n^+ \times TM_n^+, TM_n^+) \quad D \mapsto ((X, Y) \mapsto \Gamma(D)(X, Y)) \quad (2.157)$$

$$\Gamma(D)(X, Y) = \frac{1}{2} G^{(-1)}(D) \left( dG(D)(X)(Y) \right)$$

A (2.34) képlettel definiált görbületi tenzorra a

$$\begin{aligned} R(D)(X, Y)Z = & d\Gamma(D)(X)(Y, Z) + \Gamma(D) \left( X, \Gamma(D)(Y, Z) \right) - \\ & - d\Gamma(D)(Y)(X, Z) - \Gamma(D) \left( Y, \Gamma(D)(X, Z) \right) \end{aligned} \quad (2.158)$$

formulát kapjuk. A  $\Gamma(D)$  leképezés deriváltja, az inverzfüggvény valamint a függvénykompozíció deriválási szabálya alapján

$$d\Gamma(D)(X)(Y, Z) = \frac{1}{2} dG^{(-1)}(D)(X) \left( dG(D)(Y)(Z) \right) + \frac{1}{2} G^{(-1)}(D) \left( d^2 G(D)(X)(Y)(Z) \right). \quad (2.159)$$

Ennek meghatározásához szükség van a  $G$  függvény második deriváltjára valamint a  $G^{(-1)}$  függvény deriváltjára.

A  $G$  függvény második deriváltja

$$d^2G : M_n^+ \rightarrow \text{Lin}\left(TM_n^+, \text{Lin}\left(TM_n^+, \text{Lin}(TM_n^+, TM_n^+)\right)\right) \quad (2.160)$$

$$D \mapsto \left( Z \mapsto \left( Y \mapsto \left( X \mapsto d^2G(D)(Z)(Y)(X) \right) \right) \right)$$

leképezés, ahol

$$d^2G(D)(Z)(Y)(X) = \int_T (D+t)^{-1}Z(D+t)^{-1}X(D+t)^{-1}Y(D+t)^{-1} + \quad (2.161)$$

$$+ (D+t)^{-1}X(D+t)^{-1}Z(D+t)^{-1}Y(D+t)^{-1} +$$

$$+ (D+t)^{-1}X(D+t)^{-1}Y(D+t)^{-1}Z(D+t)^{-1} +$$

$$+ (D+t)^{-1}Z(D+t)^{-1}Y(D+t)^{-1}X(D+t)^{-1} +$$

$$+ (D+t)^{-1}Y(D+t)^{-1}Z(D+t)^{-1}X(D+t)^{-1} +$$

$$+ (D+t)^{-1}Y(D+t)^{-1}X(D+t)^{-1}Z(D+t)^{-1} d\mu(t) .$$

A fenti formulából látható, hogy a  $d^2G(D)(Z)(Y)(X)$  kifejezés szimmetrikus az  $X, Y$  és  $Z$  változóknban.

A  $G^{(-1)}$  függvény deriváltja

$$dG^{(-1)} : M_n^+ \rightarrow \text{Lin}\left(TM_n^+, \text{Lin}(TM_n^+, TM_n^+)\right) \quad (2.162)$$

$$D \mapsto \left( X \mapsto \left( Y \mapsto dG^{(-1)}(D)(X)(Y) \right) \right) \quad (2.163)$$

leképezés, melyre az inverzfüggvény és az összetett függvény deriválási szabálya alapján

$$dG^{(-1)}(D)(X)(Y) = -G^{(-1)}(D)\left(dG(D)(X)\left(G^{(-1)}(D)(Y)\right)\right) \quad (2.164)$$

teljesül.

A görbületi tenzor (2.158) alakjába beírva a Levi–Civita-féle kovariáns deriválásra

és a deriváltjára kapott (2.157,2.159) kifejezéseket a

$$\begin{aligned}
R(D)(X, Y)(Z) &= \frac{1}{2}dG^{(-1)}(D)(X)(dG(D)(Y)(Z)) + \frac{1}{2}G^{(-1)}(D)(d^2G(D)(X)(Y)(Z)) \\
&\quad - \frac{1}{2}dG^{(-1)}(D)(Y)(dG(D)(X)(Z)) - \frac{1}{2}G^{(-1)}(D)(d^2G(D)(Y)(X)(Z)) + \\
&\quad + \frac{1}{2}G^{(-1)}(D) \left( dG(D)(X) \left( \frac{1}{2}G^{(-1)}(D)(dG(D)(Y)(Z)) \right) \right) - \\
&\quad - \frac{1}{2}G^{(-1)}(D) \left( dG(D)(Y) \left( \frac{1}{2}G^{(-1)}(D)(dG(D)(X)(Z)) \right) \right) \quad (2.165)
\end{aligned}$$

formulát kapjuk. Kihhasználva  $d^2G(D)$  szimmetrikusságát igazoló (2.161) egyenletet, valamint a  $dG^{(-1)}$  kiszámítására kapott (2.164) kifejezést, a görbületi tenzorra

$$\begin{aligned}
R(D)(X, Y)(Z) &= \frac{1}{4}G^{(-1)}(D) \left( dG(D)(Y) \left( G^{(-1)}(D)(dG(D)(X)(Z)) \right) \right) - \quad (2.166) \\
&\quad - \frac{1}{4}G^{(-1)}(D) \left( dG(D)(X) \left( G^{(-1)}(D)(dG(D)(Y)(Z)) \right) \right)
\end{aligned}$$

adódik.

Legyen  $D \in M_n^+$  és  $(A_s)_{s \in I}$  a  $\langle \cdot, \cdot \rangle$  skalárszorzásra nézve ortonormált bázis a  $T_D M_n^+$  érintőtéren. Ekkor az  $(M_n^+, g)$  Riemann-metrika skalárgörbületére a (2.46, 2.47 2.48, 2.49) egyenletek alapján

$$\text{Scal}(D) = \sum_{t,s \in I} \left\langle R(D)(A_s, A_t) \left( G^{(-1)}(D)(A_t) \right), A_s \right\rangle \quad (2.167)$$

teljesül.

Definiáljuk az  $I_1 = \{(i, j) \mid 1 \leq i < j \leq n\}$  valamint az  $I_2 = \{(i, i) \mid 1 \leq i \leq n\}$  indexhalmazt. A  $t = (i, j) \in I_1$  esetén legyen  $A_t = \frac{F_{ij}}{\sqrt{2}}$  valamint  $t = (i, i) \in I_2$  esetén  $A_t = \frac{F_{ii}}{2}$ . Ekkor  $(A_t)_{t \in I_1 \cup I_2}$  tetszőleges  $D \in M_n^+$  pont esetén a  $T_D M_n^+$  tér ortonormált bázisa a  $\langle \cdot, \cdot \rangle$  skalárszorzásra nézve. Az  $I_1$  indexhalmazhoz tartozó  $A_t$  mátrixok fődiagonálisában csak nulla szerepel, míg az  $I_2$ -höz tartozók diagonálisak. A

$$\sum_{t,s \in I_1 \cup I_2} = \sum_{t,s \in I_1} + \sum_{t \in I_2, s \in I_1} + \sum_{s \in I_2, t \in I_1} + \sum_{t,s \in I_2} \quad (2.168)$$

felbontásnak megfelelően a (2.167) képletet négy részletben fogjuk kiszámolni. A négy összegnek az (OFF-OFF), (OFF-DIAG), (DIAG-OFF) és (DIAG-DIAG) neveket adjuk.

A (2.167) képlet első összege, (OFF-OFF):

$$\sum_{t,s \in I_1} \left\langle R(D)(A_s, A_t) \left( G^{(-1)}(D)(A_t) \right), A_s \right\rangle = \quad (2.169)$$

$$= \sum_{\substack{1 \leq i < j \leq n \\ 1 \leq k < l \leq n}} \left\langle R(D) \left( \frac{F_{kl}}{\sqrt{2}}, \frac{F_{ij}}{\sqrt{2}} \right) \frac{1}{m_{ij}} \frac{F_{ij}}{\sqrt{2}}, \frac{F_{kl}}{\sqrt{2}} \right\rangle = \quad (2.170)$$

$$= \sum_{\substack{1 \leq i < j \leq n \\ 1 \leq k < l \leq n}} \frac{1}{4m_{ij}} \langle R(D)(F_{kl}, F_{ij}) F_{ij}, F_{kl} \rangle = \quad (2.171)$$

$$= \sum_{\substack{1 \leq i < j \leq n \\ 1 \leq k < l \leq n}} \frac{-1}{16m_{ij}} \left\langle G^{(-1)}(D) \left( dG(D)(F_{kl}) \left( G^{(-1)}(D) \left( dG(D)(F_{ij})(F_{ij}) \right) \right) \right), F_{kl} \right\rangle + \quad (2.172)$$

$$+ \sum_{\substack{1 \leq i < j \leq n \\ 1 \leq k < l \leq n}} \frac{1}{16m_{ij}} \left\langle G^{(-1)}(D) \left( dG(D)(F_{ij}) \left( G^{(-1)}(D) \left( dG(D)(F_{kl})(F_{ij}) \right) \right) \right), F_{kl} \right\rangle$$

A (2.172) képlet első tagja:

$$\sum_{\substack{1 \leq i < j \leq n \\ 1 \leq k < l \leq n}} \frac{-1}{16m_{ij}} \left\langle G^{(-1)}(D) \left( dG(D)(F_{kl}) \left( G^{(-1)}(D) \left( -F_{ii} m_{ij} - m_{ij} F_{jj} \right) \right) \right), F_{kl} \right\rangle = \quad (2.173)$$

$$= \sum_{\substack{1 \leq i < j \leq n \\ 1 \leq k < l \leq n}} \frac{1}{16m_{ij}} \left\langle G^{(-1)}(D) \left( dG(D)(F_{kl}) \left( \frac{m_{iij}}{m_{ii}} F_{ii} + \frac{m_{ijj}}{m_{jj}} F_{jj} \right) \right), F_{kl} \right\rangle = \quad (2.174)$$

$$= \sum_{\substack{1 \leq i < j \leq n \\ 1 \leq k < l \leq n}} - \frac{1}{16m_{ij}} \frac{m_{iij}}{m_{ii}} \left\langle G^{(-1)}(D) (\delta_{li} m_{kli} F_{ki} + \delta_{ki} m_{ikl} F_{li} + \right. \\ \left. + \delta_{li} m_{kli} F_{ki} + \delta_{ki} m_{kli} F_{li}), F_{kl} \right\rangle + \quad (2.175)$$

$$- \frac{1}{16m_{ij}} \frac{m_{ijj}}{m_{jj}} \left\langle G^{(-1)}(D) (\delta_{lj} m_{klj} F_{kj} + \delta_{kj} m_{klj} F_{lj} + \right. \\ \left. + \delta_{lj} m_{klj} F_{kj} + \delta_{kj} m_{klj} F_{lj}), F_{kl} \right\rangle =$$

$$= - \sum_{\substack{1 \leq i < j \leq n \\ 1 \leq k < l \leq n}} \left[ \frac{1}{16m_{ij}} \frac{m_{iij}}{m_{ii}} \left( 2 \frac{m_{kli}}{m_{ki}} \delta_{li} \langle F_{ki}, F_{kl} \rangle + 2 \frac{m_{kli}}{m_{li}} \delta_{ki} \langle F_{li}, F_{kl} \rangle \right) + \right. \\ \left. + \frac{1}{16m_{ij}} \frac{m_{ijj}}{m_{jj}} \left( 2 \frac{m_{klj}}{m_{kj}} \delta_{lj} \langle F_{kj}, F_{kl} \rangle + 2 \frac{m_{klj}}{m_{lj}} \delta_{kj} \langle F_{lj}, F_{kl} \rangle \right) \right] = \quad (2.176)$$

$$= - \frac{1}{8} \sum_{1 \leq k < l \leq n} \left[ \sum_{i,j=1}^n \left( \frac{1}{m_{ij}} \frac{m_{iij}}{m_{ii}} \left( 2 \frac{m_{kli}}{m_{ki}} \delta_{li} + 2 \frac{m_{kli}}{m_{li}} \delta_{ki} \right) \right) - \right. \\ \left. - \sum_{i=1}^n \left( \frac{1}{m_{ii}} \frac{m_{iii}}{m_{ii}} \left( 2 \frac{m_{kli}}{m_{ki}} \delta_{li} + 2 \frac{m_{kli}}{m_{li}} \delta_{ki} \right) \right) \right] = \quad (2.177)$$

$$= - \frac{1}{4} \sum_{1 \leq k < l \leq n} \left[ \sum_{j=1}^n \left( \frac{1}{m_{lj}} \frac{m_{llj}}{m_{ll}} \frac{m_{klj}}{m_{kl}} + \frac{1}{m_{kj}} \frac{m_{kkj}}{m_{kk}} \frac{m_{klj}}{m_{kl}} \right) - \right. \\ \left. - \frac{1}{m_{ll}} \frac{m_{lll}}{m_{ll}} \frac{m_{klj}}{m_{kl}} - \frac{1}{m_{kk}} \frac{m_{kkk}}{m_{kk}} \frac{m_{klj}}{m_{kl}} \right] = \quad (2.178)$$

$$= - \frac{1}{4} \left[ \sum_{1 \leq k < l < j \leq n} \left( \frac{m_{llj} m_{klj}}{m_{lj} m_{ll} m_{kl}} + \frac{m_{kkj} m_{klj}}{m_{kj} m_{kk} m_{kl}} \right) + \right. \\ \left. + \sum_{1 \leq k < j < l \leq n} \left( \frac{m_{llj} m_{klj}}{m_{lj} m_{ll} m_{kl}} + \frac{m_{kkj} m_{klj}}{m_{kj} m_{kk} m_{kl}} \right) + \right. \quad (2.179)$$



$$\begin{aligned}
& + \sum_{1 \leq j < k < l \leq n} \left( \frac{m_{llj}m_{kll}}{m_{lj}m_{ll}m_{kl}} + \frac{m_{kkj}m_{kkl}}{m_{kj}m_{kk}m_{kl}} \right) + \sum_{1 \leq k < l \leq n} \left( \frac{m_{llk}m_{kll}}{m_{lk}m_{ll}m_{kl}} + \frac{m_{kkk}m_{kkl}}{m_{kk}m_{kk}m_{kl}} + \right. \\
& \left. + \frac{m_{lll}m_{kll}}{m_{ll}m_{ll}m_{kl}} + \frac{m_{kkk}m_{kkl}}{m_{kl}m_{kk}m_{kl}} \right) - \sum_{1 \leq k < l \leq n}^n \left( \frac{m_{lll}m_{kll}}{m_{ll}m_{ll}m_{kl}} + \frac{m_{kkk}m_{kkl}}{m_{kk}m_{kk}m_{kl}} \right) \Big] = \\
& = -\frac{1}{2} \left[ \sum_{1 \leq u < v < w \leq n} \left( \frac{m_{vvw}m_{uww}}{m_{vu}m_{vv}m_{vw}} + \frac{m_{uuw}m_{uww}}{m_{uu}m_{uv}m_{uw}} + \frac{m_{wvu}m_{wvw}}{m_{wu}m_{wv}m_{ww}} \right) + \right. \\
& \left. + \sum_{1 \leq k < l \leq n} \left( \frac{m_{kll}^2}{m_{kl}^2 m_{ll}} + \frac{m_{kkl}^2}{m_{kl}^2 m_{kk}} \right) \right] \tag{2.180}
\end{aligned}$$

A (2.172) képlet második tagja:

$$\begin{aligned}
& \sum_{\substack{1 \leq i < j \leq n \\ 1 \leq k < l \leq n}} \frac{1}{16m_{ij}} \left\langle G^{(-1)}(D) \left( dG(D)(F_{ij}) \left( G^{(-1)}(D) \left( dG(D)(F_{kl})(F_{ij}) \right) \right) \right), F_{kl} \right\rangle = \\
& \tag{2.181} \\
& = \sum_{\substack{1 \leq i < j \leq n \\ 1 \leq k < l \leq n}} -\frac{1}{16m_{ij}} \left\langle G^{(-1)}(D) \left( dG(D)(F_{ij}) \left( G^{(-1)}(D) \left( F_{il}m_{ijl}\delta_{jk} + F_{jk}m_{ijk}\delta_{il} + \right. \right. \right. \\
& \left. \left. \left. + F_{ik}m_{ijk}\delta_{jl} + F_{jl}m_{ijl}\delta_{ik} \right) \right) \right), F_{kl} \right\rangle = \\
& \tag{2.182} \\
& = \sum_{\substack{1 \leq i < j \leq n \\ 1 \leq k < l \leq n}} -\frac{1}{16m_{ij}} \left\langle G^{(-1)}(D) \left( dG(D)(F_{ij})(F_{il}) \frac{m_{ijl}}{m_{il}} \delta_{jk} + dG(D)(F_{ij})(F_{jk}) \frac{m_{ijk}}{m_{jk}} \delta_{il} + \right. \right. \\
& \left. \left. + dG(D)(F_{ij})(F_{ik}) \frac{m_{ijk}}{m_{ik}} \delta_{jl} + dG(D)(F_{ij})(F_{jl}) \frac{m_{ijl}}{m_{jl}} \delta_{ik} \right), F_{kl} \right\rangle = \\
& \tag{2.183}
\end{aligned}$$

$$\begin{aligned}
&= \sum_{\substack{1 \leq i < j \leq n \\ 1 \leq k < l \leq n}} \frac{1}{16m_{ij}} \left\langle G^{(-1)}(D) \left( \frac{m_{ijl}}{m_{il}} \delta_{jk} (F_{ii}m_{iij}\delta_{jl} + F_{il}m_{ijl}\delta_{ji} + F_{ji}m_{ijj}\delta_{il} + F_{jl}m_{ijl}\delta_{ii}) + \right. \right. \\
&\quad + \frac{m_{ijk}}{m_{jk}} \delta_{il} (F_{ij}m_{ijj}\delta_{jk} + F_{ik}m_{ijk}\delta_{jj} + F_{jj}m_{ijj}\delta_{ik} + F_{jk}m_{ijk}\delta_{ij}) + \\
&\quad + \frac{m_{ijk}}{m_{ik}} \delta_{jl} (F_{ii}m_{iij}\delta_{jk} + F_{ik}m_{ijk}\delta_{ji} + F_{ji}m_{ijj}\delta_{ik} + F_{jk}m_{ijk}\delta_{ii}) + \\
&\quad \left. \left. + \frac{m_{ijl}}{m_{jl}} \delta_{ik} (F_{ij}m_{ijj}\delta_{jl} + F_{il}m_{ijl}\delta_{jj} + F_{jj}m_{ijj}\delta_{il} + F_{jl}m_{ijl}\delta_{ij}) \right), F_{kl} \right\rangle = \\
\end{aligned} \tag{2.184}$$

$$\begin{aligned}
&= \frac{1}{8} \sum_{\substack{1 \leq i < j \leq n \\ 1 \leq k < l \leq n}} \frac{1}{m_{ij}} \left[ \frac{m_{ijl}m_{iij}}{m_{il}m_{ij}} \delta_{jk}\delta_{il} + \frac{m_{ijl}m_{ijj}}{m_{jl}m_{ij}} \delta_{ik}\delta_{jl} + \frac{m_{ijl}^2}{m_{il}m_{jl}} \delta_{jk} + \frac{m_{ijl}^2}{m_{il}m_{jl}} \delta_{ik} + \right. \\
&\quad + \frac{m_{ijk}m_{ijj}}{m_{jk}m_{ij}} \delta_{il}\delta_{jk} + \frac{m_{ijk}m_{iij}}{m_{ik}m_{ij}} \delta_{jl}\delta_{ik} + \frac{m_{ijk}^2}{m_{ik}m_{kj}} \delta_{il} + \frac{m_{ijk}^2}{m_{ik}m_{jk}} \delta_{jl} + \\
&\quad + \left( \frac{m_{ijl}}{m_{il}} \right)^2 \delta_{jk}\delta_{ji}\delta_{ik} + \left( \frac{m_{ijk}}{m_{jk}} \right)^2 \delta_{il}\delta_{ij}\delta_{jl} + \\
&\quad \left. + \left( \frac{m_{ijk}}{m_{ik}} \right)^2 \delta_{jl}\delta_{ji}\delta_{il} + \left( \frac{m_{ijl}}{m_{jl}} \right)^2 \delta_{ik}\delta_{ji}\delta_{jk} \right] = \\
\end{aligned} \tag{2.185}$$

$$\begin{aligned}
&= \frac{1}{8} \sum_{1 \leq k < l \leq n} \left[ \frac{m_{llk}m_{llk}}{m_{ll}m_{kl}m_{kl}} + \frac{m_{lkk}m_{lkk}}{m_{lk}m_{kk}m_{kl}} + \frac{1}{m_{kk}} \left( \frac{m_{kkl}}{m_{kl}} \right)^2 + \frac{1}{m_{ll}} \left( \frac{m_{llk}}{m_{kl}} \right)^2 + \right. \\
&\quad + \sum_{i=1}^n \left( \frac{m_{ikl}^2}{m_{il}m_{kl}m_{ik}} \right) + \sum_{j=1}^n \left( \frac{m_{jlk}^2}{m_{lk}m_{jk}m_{lj}} \right) - \left( \frac{m_{kkl}m_{kkl}}{m_{kk}m_{kl}m_{kk}} \delta_{kl} + \right. \\
&\quad \left. \left. + \frac{m_{llk}m_{lll}}{m_{ll}m_{lk}m_{ll}} \delta_{lk} + \frac{2}{m_{kk}} \left( \frac{m_{kkl}}{m_{kl}} \right)^2 + \frac{2}{m_{ll}} \left( \frac{m_{llk}}{m_{kl}} \right)^2 \right) \right] = \\
\end{aligned} \tag{2.186}$$

$$= \frac{1}{8} \sum_{k,l=1}^n \left( \frac{2m_{kkl}^2}{m_{kk}m_{kl}^2} - \frac{2m_{kkl}^2}{m_{kk}m_{kl}^2} \right) + \frac{1}{4} \sum_{1 \leq k < l \leq n} \sum_{i=1}^n \frac{m_{ikl}^2}{m_{il}m_{kl}m_{ik}} = \tag{2.187}$$

$$= \frac{1}{4} \sum_{1 \leq u < v < w \leq n} \frac{3m_{uvw}^2}{m_{uv}m_{vw}m_{wu}} + \frac{1}{4} \sum_{1 \leq k < l \leq n} \left( \frac{m_{kkl}^2}{m_{kk}m_{kl}^2} + \frac{m_{kll}^2}{m_{ll}m_{kl}^2} \right) \quad (2.188)$$

Ezzel kiszámoltuk az (OFF-OFF) tagot.

A (2.167) képlet második összege, (OFF-DIAG):

$$\sum_{t \in I_2, s \in I_1} \left\langle R(D)(A_s, A_t) \left( G^{(-1)}(D)(A_t) \right), A_s \right\rangle = \quad (2.189)$$

$$= \sum_{1 \leq k < l \leq n} \sum_{i=1}^n \left\langle R(D) \left( \frac{F_{kl}}{\sqrt{2}}, \frac{F_{ii}}{2} \right) \frac{1}{m_{ii}} \frac{F_{ii}}{2}, \frac{F_{kl}}{\sqrt{2}} \right\rangle = \quad (2.190)$$

$$= \sum_{1 \leq k < l \leq n} \sum_{i=1}^n \frac{1}{8m_{ii}} \langle R(D)(F_{kl}, F_{ii}) F_{ii}, F_{kl} \rangle = \quad (2.191)$$

$$= \sum_{1 \leq k < l \leq n} \sum_{i=1}^n \frac{-1}{32m_{ii}} \left\langle G^{(-1)}(D) \left( dG(D)(F_{kl}) \left( G^{(-1)}(D) \left( dG(D)(F_{ii})(F_{ii}) \right) \right) \right), F_{kl} \right\rangle +$$

$$+ \sum_{1 \leq k < l \leq n} \sum_{i=1}^n \frac{1}{32m_{ii}} \left\langle G^{(-1)}(D) \left( dG(D)(F_{ii}) \left( G^{(-1)}(D) \left( dG(D)(F_{kl})(F_{ii}) \right) \right) \right), F_{kl} \right\rangle \quad (2.192)$$

A (2.192) képlet első tagja:

$$\sum_{1 \leq k < l \leq n} \sum_{i=1}^n \frac{-1}{32m_{ii}} \left\langle G^{(-1)}(D) \left( dG(D)(F_{kl}) \left( G^{(-1)}(D) \left( dG(D)(F_{ii})(F_{ii}) \right) \right) \right), F_{kl} \right\rangle = \quad (2.193)$$

$$= \sum_{1 \leq k < l \leq n} \sum_{i=1}^n \frac{1}{32m_{ii}} \left\langle G^{(-1)}(D) \left( dG(D)(F_{kl}) \left( G^{(-1)}(D) (4F_{ii}m_{ii}) \right) \right), F_{kl} \right\rangle = \quad (2.194)$$

$$= \sum_{1 \leq k < l \leq n} \sum_{i=1}^n -\frac{1}{8m_{ii}} \left\langle G^{(-1)}(D) \left( \frac{m_{iii}}{m_{ii}} (2F_{ki}m_{kli}\delta_{li} + 2F_{li}m_{kli}\delta_{ki}) \right), F_{kl} \right\rangle = \quad (2.195)$$

$$= \sum_{1 \leq k < l \leq n} \sum_{i=1}^n -\frac{1}{2} \left( \frac{m_{iii} m_{kli}}{m_{ii}^2 m_{ki}} \delta_{li} + \frac{m_{iii} m_{kli}}{m_{ii}^2 m_{li}} \delta_{ki} \right) = \quad (2.196)$$

$$= -\frac{1}{2} \sum_{i=1}^n \left( \sum_{k,l=1}^n \frac{m_{kli} m_{iii}}{m_{ki} m_{ii}^2} \delta_{li} - \sum_{k=1}^n \frac{m_{kki} m_{iii}}{m_{ki} m_{ii}^2} \delta_{ki} \right) = \quad (2.197)$$

$$= -\frac{1}{2} \sum_{k,l=1}^n \frac{m_{kli} m_{kkk}}{m_{kl} m_{kk}^2} + \frac{1}{4} \sum_{k=1}^n \frac{m_{kkk}}{m_{kk}} \quad (2.198)$$

A (2.192) képlet második tagja:

$$\sum_{1 \leq k < l \leq n} \sum_{i=1}^n \frac{1}{32 m_{ii}} \left\langle G^{(-1)}(D) \left( dG(D)(F_{ii}) \left( G^{(-1)}(D) \left( dG(D)(F_{kl})(F_{ii}) \right) \right) \right), F_{kl} \right\rangle = \quad (2.199)$$

$$= \sum_{1 \leq k < l \leq n} \sum_{i=1}^n -\frac{1}{16 m_{ii}} \left\langle G^{(-1)}(D) \left( \frac{m_{kli}}{m_{ki}} \delta_{il} dG(D)(F_{ii})(F_{ki}) + \right. \right. \quad (2.200)$$

$$\left. \left. + \frac{m_{kli}}{m_{li}} \delta_{ik} dG(D)(F_{ii})(F_{li}) \right), F_{kl} \right\rangle =$$

$$= \frac{1}{4} \sum_{1 \leq k < l \leq n} \sum_{i=1}^n \frac{1}{m_{ii}} \left( \frac{m_{kli} m_{iik}}{m_{ik}^2} \delta_{il} + \frac{m_{kli} m_{iil}}{m_{il}^2} \delta_{ik} + 2 \frac{m_{kli} m_{iii}}{m_{ik} m_{ii}} \delta_{ik} \delta_{il} + 2 \frac{m_{kli} m_{iii}}{m_{il} m_{ii}} \delta_{ik} \delta_{il} \right) = \quad (2.201)$$

$$= \frac{1}{4} \sum_{i=1}^n \left[ \sum_{k,l=1}^n \left( \frac{m_{kli} m_{iik}}{m_{ii} m_{ki}^2} \delta_{il} + 2 \frac{m_{kli} m_{iii}}{m_{ik} m_{ii}^2} \delta_{ik} \delta_{il} \right) - \right. \quad (2.202)$$

$$\left. - \sum_{k=1}^n \left( \frac{m_{kki} m_{iik}}{m_{ii} m_{ki}^2} \delta_{ik} + 2 \frac{m_{kki} m_{iii}}{m_{ik} m_{ii}^2} \delta_{ik} \right) \right] =$$

$$= \frac{1}{4} \sum_{i=1}^n \left[ \sum_{k=1}^n \frac{m_{iik}^2}{m_{ii} m_{ik}^2} - \frac{m_{iii}^2}{m_{ii}^3} \right] = \quad (2.203)$$

$$= \frac{1}{4} \sum_{k,l=1}^n \frac{m_{kkl}^2}{m_{kk} m_{kl}^2} - \frac{1}{4} \sum_{k=1}^n \frac{m_{kkk}^2}{m_{kk}^3} \quad (2.204)$$

Ezzel kiszámoltuk az (OFF-DIAG) tagot.

A (2.167) képlet harmadik összege, (DIAG-OFF):

$$\sum_{t \in I_1, s \in I_2} \left\langle R(D)(A_s, A_t) \left( G^{(-1)}(D)(A_t) \right), A_s \right\rangle = \quad (2.205)$$

$$= \sum_{1 \leq k < l \leq n} \sum_{i=1}^n \left\langle R(D) \left( \frac{F_{ii}}{2}, \frac{F_{kl}}{\sqrt{2}} \right) \frac{1}{m_{kl}} \frac{F_{kl}}{\sqrt{2}}, \frac{F_{ii}}{2} \right\rangle = \quad (2.206)$$

$$= \sum_{1 \leq k < l \leq n} \sum_{i=1}^n \frac{1}{8m_{kl}} \langle R(D)(F_{ii}, F_{kl}) F_{kl}, F_{ii} \rangle = \quad (2.207)$$

$$= \sum_{1 \leq k < l \leq n} \sum_{i=1}^n \frac{-1}{32m_{kl}} \left\langle G^{(-1)}(D) \left( dG(D)(F_{ii}) \left( G^{(-1)}(D) \left( dG(D)(F_{kl})(F_{kl}) \right) \right) \right), F_{ii} \right\rangle +$$

$$+ \sum_{1 \leq k < l \leq n} \sum_{i=1}^n \frac{1}{32m_{kl}} \left\langle G^{(-1)}(D) \left( dG(D)(F_{kl}) \left( G^{(-1)}(D) \left( dG(D)(F_{ii})(F_{kl}) \right) \right) \right), F_{ii} \right\rangle \quad (2.208)$$

A (2.208) képlet első tagja:

$$\sum_{1 \leq k < l \leq n} \sum_{i=1}^n \frac{-1}{32m_{kl}} \left\langle G^{(-1)}(D) \left( dG(D)(F_{ii}) \left( G^{(-1)}(D) \left( dG(D)(F_{kl})(F_{kl}) \right) \right) \right), F_{ii} \right\rangle = \quad (2.209)$$

$$= \sum_{1 \leq k < l \leq n} \sum_{i=1}^n \frac{1}{32m_{kl}} \left\langle G^{(-1)}(D) \left( dG(D)(F_{ii}) \left( G^{(-1)}(D) (F_{kk}m_{kkl} + F_{ll}m_{kll}) \right) \right), F_{ii} \right\rangle = \quad (2.210)$$

$$= \sum_{1 \leq k < l \leq n} \sum_{i=1}^n \frac{1}{32m_{kl}} \left\langle G^{(-1)}(D) \left( dG(D)(F_{ii}) \left( \frac{m_{kkl}}{m_{kk}} F_{kk} + \frac{m_{kll}}{m_{ll}} F_{ll} \right) \right), F_{ii} \right\rangle = \quad (2.211)$$

$$= \sum_{1 \leq k < l \leq n} \sum_{i=1}^n -\frac{1}{8m_{kl}} \left\langle \frac{m_{kkl}m_{kii}}{m_{kk}m_{ki}} \delta_{ki} F_{ik} + \frac{m_{kll}m_{lii}}{m_{ll}m_{li}} \delta_{li} F_{il}, F_{ii} \right\rangle = \quad (2.212)$$

$$= \sum_{1 \leq k < l \leq n} -\frac{1}{2m_{kl}} \left( \frac{m_{kk}m_{kkk}}{m_{kk}^2} + \frac{m_{kl}m_{lll}}{m_{ll}^2} \right) \quad (2.213)$$

Ez megegyezik az OFF-DIAG első tagjával, a (2.196) képlettel. Vagyis ez a tag megegyezik a (2.198) képlettel.

A (2.208) képlet második tagja:

$$\sum_{1 \leq k < l \leq n} \sum_{i=1}^n \frac{1}{32m_{kl}} \left\langle G^{(-1)}(D) \left( dG(D)(F_{kl}) \left( G^{(-1)}(D) \left( dG(D)(F_{ii})(F_{kl}) \right) \right) \right), F_{ii} \right\rangle = \quad (2.214)$$

$$= \sum_{1 \leq k < l \leq n} \sum_{i=1}^n -\frac{1}{16m_{kl}} \left\langle G^{(-1)}(D) \left( dG(D)(F_{kl}) \left( G^{(-1)}(D) \left( \frac{m_{lii}}{m_{li}} \delta_{ik} F_{il} + \right. \right. \right. \right. \quad (2.215)$$

$$\left. \left. \left. \left. + \frac{m_{kii}}{m_{ki}} \delta_{il} F_{ik} \right) \right) \right), F_{ii} \right\rangle =$$

$$= \sum_{1 \leq k < l \leq n} \sum_{i=1}^n \frac{1}{16m_{kl}} \left\langle G^{(-1)}(D) \left( \frac{m_{lii}}{m_{li}} \delta_{ik} (F_{kl}m_{kll}\delta_{il} + F_{ki}m_{kli} + F_{li}m_{kll}\delta_{ik}) + \right. \right. \quad (2.216)$$

$$\left. \left. + \frac{m_{kii}}{m_{ki}} \delta_{il} (F_{kk}m_{kkl}\delta_{il} + F_{li}m_{kli} + F_{lk}m_{kkl}\delta_{il}) \right), F_{ii} \right\rangle =$$

$$= \sum_{1 \leq k < l \leq n} \sum_{i=1}^n \frac{1}{16m_{kl}} \left( 4 \frac{m_{lii}m_{kli}}{m_{il}m_{ik}} \delta_{ik} + 4 \frac{m_{kii}m_{kli}}{m_{ik}m_{il}} \delta_{il} \right) = \quad (2.217)$$

$$= \sum_{1 \leq k < l \leq n} \frac{1}{4m_{kl}} \left( \frac{m_{kkl}^2}{m_{kl}m_{kk}} + \frac{m_{kll}^2}{m_{kl}m_{ll}} \right) = \quad (2.218)$$

$$= \frac{1}{4} \sum_{k,l=1}^n \frac{m_{kkl}^2}{m_{kl}^2 m_{kk}} - \frac{1}{4} \sum_{k=1}^n \frac{m_{kkk}^2}{m_{kk}^3} \quad (2.219)$$

Ez megegyezik az OFF-DIAG második tagjával, a (2.204) egyenlettel.

Ezzel kiszámoltuk a (DIAG-OFF) tagot.

A (2.167) képlet negyedik összege, (DIAG-DIAG):

$$\sum_{t,s \in I_2} \left\langle R(D)(A_s, A_t) \left( G^{(-1)}(D)(A_t) \right), A_s \right\rangle = \quad (2.220)$$

$$= \sum_{i,j=1}^n \left\langle R(D) \left( \frac{F_{ii}}{2}, \frac{F_{jj}}{2} \right) \frac{1}{m_{jj}} \frac{F_{jj}}{2}, \frac{F_{ii}}{2} \right\rangle = \quad (2.221)$$

$$= \sum_{i,j=1}^n \frac{1}{16m_{jj}} \langle R(D)(F_{ii}, F_{jj}) F_{jj}, F_{ii} \rangle = \quad (2.222)$$

$$= \sum_{i,j=1}^n \frac{1}{64m_{jj}} \left\langle G^{(-1)}(D) \left( dG(D)(F_{jj}) \left( G^{(-1)}(D) \left( dG(D)(F_{ii})(F_{jj}) \right) \right) \right) - \right. \\ \left. - G^{(-1)}(D) \left( dG(D)(F_{ii}) \left( G^{(-1)}(D) \left( dG(D)(F_{jj})(F_{jj}) \right) \right) \right), F_{ii} \right\rangle = \quad (2.223)$$

$$= -\frac{1}{64} \sum_{i,j=1}^n \frac{1}{m_{jj}} \left\langle G^{(-1)}(D) \left( 4 \frac{m_{ij}}{m_{ij}} \delta_{ij} dG(D)(F_{jj})(F_{ij}) - \right. \right. \quad (2.224)$$

$$\left. - 4 \frac{m_{jj}}{m_{jj}} dG(D)(F_{ii})(F_{jj}) \right), F_{ii} \right\rangle =$$

$$= \frac{1}{16} \sum_{i,j=1}^n \frac{1}{m_{jj}} \left\langle 2 \frac{m_{ij} m_{jji}}{m_{ij}^2} \delta_{ij} F_{ij} + 2 \frac{m_{ij} m_{jjj}}{m_{ij} m_{jj}} \delta_{ij} F_{jj} - 4 \frac{m_{ij} m_{jjj}}{m_{ij} m_{jj}} \delta_{ij} F_{ij}, F_{ii} \right\rangle = \quad (2.225)$$

$$= \frac{1}{16} \sum_{i=1}^n \frac{1}{m_{ii}} \left( 8 \left( \frac{m_{iii}}{m_{ii}} \right)^2 + 8 \left( \frac{m_{iii}}{m_{ii}} \right)^2 - 16 \left( \frac{m_{iii}}{m_{ii}} \right)^2 \right) = 0 \quad (2.226)$$

Ezzel kiszámoltuk a (DIAG-DIAG) tagot.

A skalárgörbület kiszámításához az alábbi tagokat kell figyelembe venni a (2.168) képletnek megfelelően:

1-szer: OFF-OFF, a (2.180) és a (2.188) képletek összege;

2-szer: OFF-DIAG, a (2.198) és a (2.204) képletek összege;

1-szer: DIAG-DIAG, a (2.226) képlet.

Ezek alapján a skalárgörbületre

$$\begin{aligned} \text{Scal}_D^{\mathbb{R}} = & \frac{1}{4} \sum_{1 \leq k < l \leq n} \left[ \sum_{j=1}^n \left( \frac{m_{klj}^2}{m_{kl}m_{lj}m_{jk}} - \frac{m_{kk}m_{kkj}}{m_{kj}m_{kk}m_{kl}} - \frac{m_{llj}m_{llk}}{m_{lj}m_{ll}m_{lk}} \right) \right] + \\ & + \frac{1}{8} \sum_{1 \leq k < l \leq n} \frac{m_{kkl} + m_{kll}}{m_{kl}} + \frac{1}{2} \sum_{k,l=1}^n \left( \frac{m_{kk}^2}{m_{kk}m_{kl}^2} - \frac{m_{kkl}}{m_{kl}} \right) + \frac{1}{8} \sum_{k=1}^n m_{kk} \end{aligned} \quad (2.227)$$

teljesül. Ennek átalakításával kapjuk a

$$\begin{aligned} \text{Scal}_D^{\mathbb{R}} = & \frac{1}{8} \sum_{j,k,l=1}^n \frac{m_{jkl}^2}{m_{jk}m_{kl}m_{lj}} - \frac{1}{8} \sum_{k,l=1}^n \frac{m_{kk}^2}{m_{kk}m_{kl}^2} - \frac{1}{4} \sum_{j,k,l=1}^n \frac{m_{kk}m_{kkj}}{m_{kk}m_{kl}m_{kj}} + \frac{1}{8} \sum_{k,l=1}^n \frac{m_{kkl}}{m_{kl}} + \\ & + \frac{1}{2} \sum_{k,l=1}^n \left( \frac{m_{kk}^2}{m_{kk}m_{kl}^2} - \frac{m_{kkl}}{m_{kl}} \right) - \frac{1}{8} \sum_{k=1}^n m_{kk} + \frac{1}{8} \sum_{k,l=1}^n \frac{m_{kkl}}{m_{kl}} = \\ = & \frac{1}{4} \sum_{j,k,l=1}^n \left( \frac{1}{2} \frac{m_{jkl}^2}{m_{jk}m_{kl}m_{lj}} - \frac{m_{kk}m_{kkj}}{m_{kk}m_{kl}m_{kj}} \right) + \frac{1}{4} \sum_{k,l=1}^n \left( \frac{1}{2} \frac{m_{kk}^2}{m_{kl}^2m_{kk}} - \frac{m_{kkl}m_{kll}}{m_{kl}^2m_{ll}} \right) + \\ & + \frac{1}{16} \sum_{k=1}^n m_{kk} \end{aligned} \quad (2.228)$$

kifejezéseket. Ezek alapján

$$\text{Scal}_D^{\mathbb{R}} = \frac{1}{4} \sum_{\substack{j,k,l=1 \\ |\{j,k,l\}| > 1}}^n \left( \frac{1}{2} \frac{m_{jkl}^2}{m_{jk}m_{kl}m_{lj}} - \frac{m_{kk}m_{kkj}}{m_{kk}m_{kl}m_{kj}} \right) + \frac{1}{4} \sum_{\substack{k,l=1 \\ k \neq l}}^n \left( \frac{1}{2} \frac{m_{kk}^2}{m_{kl}^2m_{kk}} - \frac{m_{kkl}m_{kll}}{m_{kl}^2m_{ll}} \right) \quad (2.229)$$

teljesül, ahol  $|\{j, k, l\}| > 1$  jelentése, hogy a  $\{j, k, l\}$  halmaznak legalább két eleme van, vagyis a  $j, k, l$  indexek nem mind azonosak.

Visszatérve a normális eloszlás esetére, vizsgáljuk meg, hogy a bevezetett differenciálgeometria mennyiségeket hogyan lehet meghatározni.



A (2.124) pontban bevezetett  $G$  leképezés, ekkor az

$$G(D)(X) = D^{-1}XD^{-1} \quad (2.230)$$

alakot ölti. Ennek a deriváltja, az (2.143) egyenlet alapján

$$dG(D)(X)(Y) = -D^{-1}XD^{-1}YD^{-1} - D^{-1}YD^{-1}XD^{-1} \quad (2.231)$$

lesz. A Levi–Civita-féle kovariáns deriválásra a (2.157) egyenlet alapján

$$\Gamma(D)(X, Y) = -\frac{1}{2}(XD^{-1}Y + YD^{-1}X) \quad (2.232)$$

adódik. A görbületi tenzor a (2.166) képlet alapján így alakul.

$$\begin{aligned} R(D)(X, Y)(Z) &= \frac{1}{4}D \left( D^{-1}YD^{-1} \left( D(D^{-1}XD^{-1}ZD^{-1} + D^{-1}ZD^{-1}XD^{-1})D \right) D^{-1} + \right. \\ &\quad \left. + D^{-1} \left( D(D^{-1}XD^{-1}ZD^{-1} + D^{-1}ZD^{-1}XD^{-1})D \right) D^{-1}YD^{-1} \right) D - \\ &\quad - \frac{1}{4}D \left( D^{-1}XD^{-1} \left( D(D^{-1}YD^{-1}ZD^{-1} + D^{-1}ZD^{-1}YD^{-1})D \right) D^{-1} + \right. \\ &\quad \left. + D^{-1} \left( D(D^{-1}YD^{-1}ZD^{-1} + D^{-1}ZD^{-1}YD^{-1})D \right) D^{-1}YD^{-1} \right) D = \\ &= \frac{1}{4}(YD^{-1}XD^{-1}Z + ZD^{-1}XD^{-1}Y - XD^{-1}YD^{-1}Z - ZD^{-1}YD^{-1}X) \end{aligned} \quad (2.233)$$

A skalárgörbületet a (2.167) egyenlet alapján a

$$\begin{aligned} \text{Scal}(D) &= \frac{1}{4} \sum_{t,s \in I} \text{Tr} (A_t D^{-1} A_s A_t D A_s + D A_t A_s D^{-1} A_t A_s - \\ &\quad - A_s D^{-1} A_t A_t D A_s - D A_t A_t D^{-1} A_s A_s) \end{aligned} \quad (2.234)$$

képlettel lehet kiszámolni, melyet a

$$\text{Scal}(D) = \frac{1}{4} \sum_{t,s \in I} \text{Tr} [A_t D^{-1}, A_s D^{-1}] [D A_t, D A_s] \quad (2.235)$$

formában is felírhatunk.

Az  $m_{ij} = (2\lambda_i\lambda_j)^{-1}$  és az  $m_{ijk} = (2\lambda_i\lambda_j\lambda_k)^{-1}$  egyenlőségek alapján a skalárgörbületre az alábbi kifejezést kapjuk.

$$\text{Scal} = -\frac{n(n-1)(n+2)}{16} \quad (2.236)$$

Tehát az  $n$ -dimenziós normális eloszlásnál a Fisher-féle metrika esetén a paraméterter skálárgörbülete állandó.

## 2.7. Az eloszlások geometriai távolsága

A valószínűségieloszlások közötti távolság mérésének módjait az (1.3) fejezetben vizsgáltuk. Az általánosított divergenciák voltak a távolságfüggvények. Az (1.8.) tételből kiderült, hogy a legtöbbet használt divergenciák második deriváltjai mind a Fisher-féle információ pozitív számszorosát adják. A Fisher-féle információval Riemann-sokasággá tehetjük a statisztikai modellünk paraméterterét. Ebben a térben a pontok eloszlásoknak felelnek meg. Adott két pont esetén az őket összekötő görbének a hosszát kiszámolhatjuk a (2.51) képlet segítségével. Ezzel a módszerrel definiálhatjuk két eloszlásnak az őket összekötő úttól függő távolságát. Természetes távolságfogalmat kapunk, ha a két pont távolságát az őket összekötő legrövidebb út hosszaként definiáljuk.

**2.22. Definíció.** Legyen  $(X, S, \Xi)$  statisztikai modell,  $p, q \in S$  és jelölje  $g^{(F)}$  a Fisher-féle információból származó Riemann-metrikát. Ekkor a  $p$  és  $q$  pont *távolsága* legyen

$$d(p, q) = \inf \left\{ \int_0^1 \sqrt{g^{(F)}(\dot{\gamma}(t), \dot{\gamma}(t))} \, dt \quad \left| \quad \begin{array}{l} \gamma : [0, 1] \rightarrow S : \gamma(0) = p, \gamma(1) = q, \\ \gamma \text{ szakaszonként } C^\infty\text{-beli} \end{array} \right. \right\}. \quad (2.237)$$

A (2.2) fejezetben említettük, hogy a Riemann-sokaságon a geodetikusok felelnek meg az egyeneseknek. A  $p, q \in S$  eloszlások közötti távolságot definiálhatnánk az őket összekötő geodetikus hosszával, azonban elképzelhető, hogy több olyan geodetikus van, mely  $p$ -ból  $q$ -ba halad. Ha a  $p$  és  $q$  pontok elég közel vannak egymáshoz, akkor már van egy kitüntetett geodetikus, mely összeköti őket.

**2.25. Tétel.** Legyen  $(M, g)$  Riemann-sokaság és  $p \in M$  tetszőleges pont. Ekkor létezik egy  $U_p$  környezete a  $p$  pontnak, hogy minden  $q \in U_p$  esetén pontosan egy  $\gamma : [0, 1] \rightarrow M$  olyan geodetikus van, melyre  $\gamma(0) = p$ ,  $\gamma(1) = q$  és  $\text{Ran } \gamma \subseteq U_p$  teljesül.

Ezek alapján megvizsgáljuk, hogy különböző statisztikai modellek esetén mi lesz az eloszlások közötti  $d(\cdot, \cdot)$  távolságfüggvény.

**2.5. Példa.** Diszkrét eloszlás: A geodetikusok egyenletét már meghatároztuk a  $(\mathcal{P}_1, g^{(F)})$  és  $(\mathcal{P}_2, g^{(F)})$  terekben ((2.10.) és a (2.12.) tételben). Az általános  $(\mathcal{P}_n, g^{(F)})$  esetben is megadható a geodetikus a következő észrevételek alapján.

Tekintsük az

$$f : \mathcal{P}_n \rightarrow \mathbb{R}^{n+1} \quad (\vartheta_1, \dots, \vartheta_n) \mapsto (\sqrt{\vartheta_1}, \dots, \sqrt{\vartheta_n}, \sqrt{1 - \vartheta_1 - \dots - \vartheta_n}) \quad (2.238)$$

leképezést, mely diffeomorfizmus a  $\mathcal{P}_n$  sokaság és az  $\mathbb{R}^{n+1}$  térben lévő egységgömb pozitív ortánsba eső része között. Az  $\mathbb{R}^{n+1}$  térben az euklideszi metrika indukál egy Riemann-metrikát az egységgömbön, melynek az  $f$  általi visszahúzottja megegyezik  $g^{(F)}$ -fel.

A gömbön a geodetikusok a főkörök lesznek és a pozitív ortánsba eső részen két pontot pontosan egy olyan geodetikus köt össze, mely nem hagyja el a pozitív ortánst.

Ezek alapján a  $p, q \in S$ ,  $p = (p_1, \dots, p_n)$ ,  $q = (q_1, \dots, q_n)$  pontokat összekötő

$$\gamma : [0, t_0] \rightarrow \mathcal{P}_n \quad t \mapsto (\gamma_1(t), \dots, \gamma_n(t)) \quad (2.239)$$

geodetikus egyenlete

$$\forall i \in \{1, \dots, n\} : \quad \gamma_i(t) = \left( \frac{\sqrt{q_i} - \sqrt{p_i} \sum_{k=0}^n \sqrt{p_k q_k}}{\sqrt{1 - (\sum_{k=0}^n \sqrt{p_k q_k})^2}} \sin t + \sqrt{p_i} \cos t \right)^2, \quad (2.240)$$

ahol  $p_0 = 1 - p_1 - \dots - p_n$  és  $q_0 = 1 - q_1 - \dots - q_n$ . A  $t_0$  paraméter pedig

$$t_0 = \arccos \left( \sum_{k=0}^n \sqrt{p_k q_k} \right). \quad (2.241)$$

A  $\gamma$  geodetikusra a

$$\gamma(0) = (p_1, \dots, p_n), \quad \gamma(t_0) = (q_1, \dots, q_n) \quad (2.242)$$

feltételek teljesülnek. Ezek alapján a  $p$  és  $q$  pontok távolsága

$$d(p, q) = \arccos \left( \sum_{k=0}^n \sqrt{p_k q_k} \right). \quad (2.243)$$

Ez a távolságfüggvény divergencia, és az irodalomban olykor szintén Bhattacharyya-féle divergenciának nevezik. (Bhattacharyya-féle divergenciának mi az (1.94) képlet által definiáltat nevezzük.) □

**2.6. Példa.** Speciális többdimenziós normális eloszlás: Legyen valamilyen  $n$  természetes számra  $X = \mathbb{R}^n$ ,  $\Xi := (\mathbb{R}^+)^n$  és az  $S$  halmaz elemei legyenek a

$$p : X \times \Xi \rightarrow \mathbb{R} \quad (x_1, \dots, x_n, \vartheta_1, \dots, \vartheta_n) \mapsto p(x_1, \dots, x_n, \vartheta_1, \dots, \vartheta_n), \quad (2.244)$$

$$p(x_1, \dots, x_n, \vartheta_1, \dots, \vartheta_n) = \frac{1}{\sqrt{(2\pi)^n \prod_{k=1}^n \vartheta_k}} \exp\left(-\frac{1}{2} \sum_{k=1}^n \frac{x_k^2}{\vartheta_k}\right)$$

paraméteres sűrűségfüggvények. Tekintsük az  $(X, S, \Xi)$  statisztikai modellt. A Fisher-féle információs mátrix ekkor

$$g_{ik}^{(F)} = \delta_{ik} \frac{2}{\vartheta_i^2}. \quad (2.245)$$

A  $\gamma : \mathbb{R} \rightarrow \Xi$  geodetikus differenciálegyenletrendszer

$$\forall i \in \{1, \dots, n\} : \quad \frac{d^2 \gamma_i(t)}{dt^2} - \frac{1}{\gamma_i(t)} \left( \frac{d \gamma_i(t)}{dt} \right)^2 = 0, \quad (2.246)$$

melynek a  $\alpha_i, \beta_i$  paramétereket tartalmazó általános megoldása

$$\forall i \in \{1, \dots, n\} : \quad \gamma_i(t) = \alpha_i e^{\beta_i t}. \quad (2.247)$$

Ezek alapján a  $(\sigma_1^{(1)}, \dots, \sigma_n^{(1)}), (\sigma_1^{(2)}, \dots, \sigma_n^{(2)}) \in \Xi$  pontok geodetikus távolsága

$$d\left((\sigma_1^{(1)}, \dots, \sigma_n^{(1)}), (\sigma_1^{(2)}, \dots, \sigma_n^{(2)})\right) = \sqrt{2} \sqrt{\sum_{k=1}^n \left( \log \frac{\sigma_k^{(1)}}{\sigma_k^{(2)}} \right)^2}. \quad (2.248)$$

□

**2.7. Példa.** Többdimenziós normális eloszlás: Tekintsük a többdimenziós normális eloszlás alábbi paraméterezését. Adott  $n$  természetes szám esetén az alaphalmaz  $X = \mathbb{R}^n$ , a paraméterter a valós  $n \times n$ -es önadjungált pozitív mátrixok halmaza, vagyis

$$\tilde{M}_n = \{D \in M(\mathbb{R}, n) \mid D = D^*, D > 0\}. \quad (2.249)$$

Az  $S$  halmaz elemei legyenek az

$$g : \tilde{M}_n \times X \rightarrow \mathbb{R} \quad (D, \underline{x}) \mapsto g(D, \underline{x}) = \frac{1}{\sqrt{(2\pi)^n \det D}} \exp\left(-\frac{1}{2} \langle \underline{x}, D^{-1} \underline{x} \rangle\right) \quad (2.250)$$

paraméteres sűrűségfüggvények. Tekintsük az  $(X, S, \tilde{M}_n)$  statisztikai modellt.

A (2.23.) tételhez hasonlóan igazolható, hogy az  $(X, S, \tilde{M}_n)$  statisztikai modell esetén a Fisher-metrikára minden  $D \in \tilde{M}_n$  pont és  $X, Y \in T_D \tilde{M}_n$  vektorok esetén

$$g^{(F)}(D)(X, Y) = \int_{\mathbb{R}^n} \frac{1}{g(D, \underline{x})} \frac{\partial g(D, \underline{x})}{\partial X} \frac{\partial g(D, \underline{x})}{\partial Y} d\underline{x} = \frac{1}{2} \text{Tr}(D^{-1} X D^{-1} Y) \quad (2.251)$$

teljesül.

Legyen  $D_1, D_2 \in \tilde{M}_n$ , melyre  $D_1 D_2 = D_2 D_1$  teljesül. Definiáljuk az  $A = D_1$  és a  $B = \ln(D_2 D_1^{-1})$  mátrixokat valamint a

$$\gamma : [0, 1] \rightarrow M_n^+ \quad t \mapsto \gamma(t) = A e^{tB} \quad (2.252)$$

függvényt. Ekkor

$$\gamma(0) = D_1 \quad \gamma(1) = D_2 \quad \dot{\gamma}(t) = A e^{tB} B \quad (2.253)$$

teljesül. A  $\gamma$  görbe geodetikus, ugyanis minden  $t \in [0, 1]$  esetén

$$\ddot{\gamma}(t) + \Gamma(\gamma(t))(\dot{\gamma}(t))(\dot{\gamma}(t)) = \ddot{\gamma}(t) + \frac{1}{2} G^{(-1)}(\gamma(t)) \left( dG(\gamma(t))(\dot{\gamma}(t))(\dot{\gamma}(t)) \right) \quad (2.254)$$

$$\begin{aligned} &= \ddot{\gamma}(t) - \gamma(t) \gamma(t)^{-1} \dot{\gamma}(t) \gamma(t)^{-1} \dot{\gamma}(t) \gamma(t)^{-1} \gamma(t) \\ &= A e^{tB} B^2 - A e^{tB} B e^{-Bt} A^{-1} A e^{tB} B = 0 \end{aligned} \quad (2.255)$$

teljesül.

A  $D_1$  és a  $D_2$  pontok távolsága

$$d(D_1, D_2) = \int_0^1 \sqrt{g^{(F)}(\gamma(t))(\dot{\gamma}(t), \dot{\gamma}(t))} dt = \frac{1}{\sqrt{2}} \sqrt{\text{Tr} B^2} \quad (2.256)$$

ami másképp

$$d(D_1, D_2) = \frac{1}{\sqrt{2}} \sqrt{\text{Tr}(\ln(D_2 D_1^{-1}))^2}. \quad (2.257)$$

Diagonális  $D_1$  és  $D_2$  mátrixok esetén ez a képlet visszaadja a (2.248) egyenletet.

Ha

$$D_1 = \begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix} \quad D_2 = \begin{pmatrix} a & c \\ c & a \end{pmatrix}, \quad (2.258)$$

akkor

$$d(D_1, D_2) = \frac{1}{\sqrt{2}} \sqrt{\left[ \ln \left( 1 - \frac{c}{a} \right) \right]^2 + \left[ \ln \left( 1 + \frac{c}{a} \right) \right]^2} \quad (2.259)$$

teljesül. □



### 3. Kvantum-információgeometria

Az előző fejezetekben a klasszikus statisztika bizonyos részeit, illetve annak differenciálgeometriai módszereit tekintettük át. A jelen fejezet célja az eddigiek kiterjesztése a nemkommutatív esetre. Ez a kiterjesztés messze nem öncélú. Bizonyos kvantummechanikai modellek szerint ezen matematikai kiterjesztések hasznosak lehetnek a fizikában; annyi azonban már biztos, hogy a matematika egyes területein komoly eredményeket értek el ezen eszközökkel.

A statisztika és a statisztika differenciálgeometriai eszközeinek az ilyen irányú kiterjesztésével szórványos esetektől eltekintve, az 1990-es évektől próbálkoztak. Mint a legtöbb esetben, itt is nehéz elsőre megtalálni az általánosítás helyes módját, ami azonban utólag sokszor már „természetesnek” tűnik. Az új témakör bizonyos alapfogalmai mára már letisztázódtak, azonban sok esetben még nincs általánosan elfogadott megközelítési mód. Különösen a duális geometriák definiálása terén meglehetősen sokféle elképzeléssel lehet találkozni (azonban egyéb esetekben is vannak még eltérések). A nem teljesen letisztázott esetekben a legáltalánosabb, a legtöbb különböző nézetet magába foglaló megközelítést mutatjuk be.

Az első részben a nemkommutatív kiterjesztés fizikai eredetét tekintjük át. Ez a rész nem tartozik szorosan a kvantum-információgeometria matematikai bemutatásához, sem a dolgozat fő irányvonalához; pusztán a teljesség kedvéért említjük meg, hogy mely kvantummechanikai modellben jelennek meg az információgeometria főbb matematikai objektumai.

A második részben definiáljuk a kvantummechanikai formalizmus főbb építőelemei és részletesen megnézzük, hogy a klasszikus  $\mathcal{P}_2$  illetve  $\mathcal{P}_3$  statisztikai modell nemkommutatív általánosításai mennyire térnek el a klasszikus esettől.

A harmadik részben az eloszlás rendezetlenségét jellemző entrópiafüggvényt és az állapotok közötti majorizációs relációt általánosítjuk. Az ezen fogalmakkal kapcsolatos főbb tételek a jelen esetben is érvényben maradnak, azonban ettől a ponttól már bonyolultabbá válik az általánosítás menete.

A negyedik részben részletesen megnézzük, hogy mennyire nem triviális a Fisher-féle információ általánosítása (még egyparaméteres statisztikai sokaság esetén sem). Ennek megfelelően definiálunk néhány ésszerűnek tűnő Fisher-féle információ mennyiséget egyparaméteres statisztikai sokaságok esetére, és megvizsgáljuk a nemkommutatív Fisher-információ meghatározásának lehetőségeit. A rész végén bemutatjuk a probléma teljes megoldását jelentő Petz-tételt, a hozzá szükséges fogalmak ismertetése után. Ezen tétel szerint a nemkommutatív esetben a Fisher-féle információ mennyiségek bizonyos operátormonoton függvényekkel indexelhetők.

Az ötödik részben ezeket az operátormonoton függvényekkel indexelhető Fisher-féle információs mennyiségeket, illetve az általuk meghatározott metrikát vizsgáljuk meg. Az általános jellemzésen kívül részletesebben foglalkozunk az irodalomban is nagyobb figyelmet kapó speciális esetekkel. Továbbá bemutatjuk a klasszikus Cramer–Rao-tétel általánosítását, és megnézzük, hogy a nemkommutatív entrópia második deriváltja melyik Fisher-féle információt generálja.

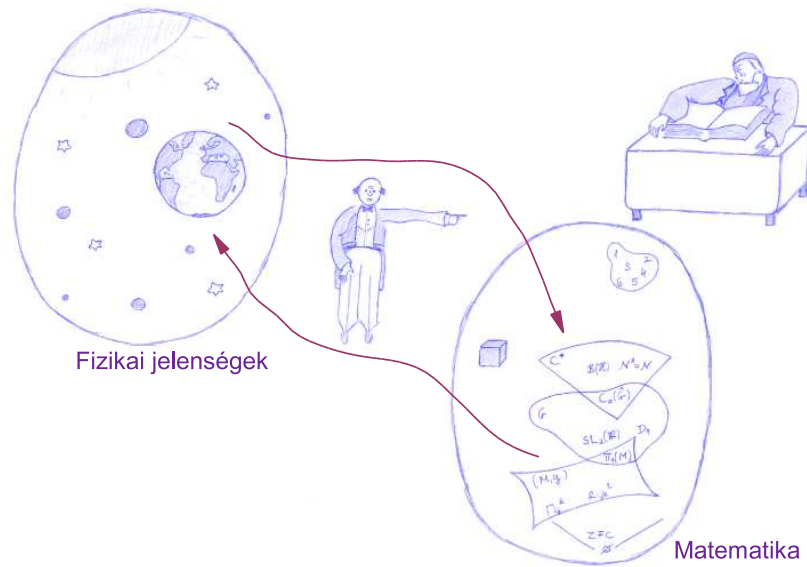
A klasszikus esetben a sokat használt relatív entrópiák szoros kapcsolatban voltak a Fisher-információval. A nemkommutatív esetben ez a kapcsolat bonyolultabb, ugyanis már a relatív entrópiák definiálásakor is komoly nehézségekbe lehet ütközni. A Csiszár-féle relatív entrópia általánosításának a segítségével azonban ismét elmondhatjuk, hogy minden relatív entrópia egy Fisher-féle információt fog definiálni, és minden Fisher-féle információ előállítható ilyen módon. A hatodik részben ezek mellett még példákön keresztül részletesen is megvizsgáljuk néhány esetben a relatív entrópiák és a Fisher-féle információk közötti kapcsolatot.

### 3.1. A kvantummechanikai modell

A később vizsgálandó matematikai objektumok fontos szerepet játszanak a kvantummechanikában. Ezen szerep pontos megfogalmazásához bizonyos előfeltevésekkel kell élnünk a fizikai világ és a matematikai világ kapcsolatáról. Ezen kapcsolat elemzése azonban nem csak a matematika tárgya. Mélyreható filozófiai elmélkedés helyett egy egyszerűsített elképzelést mutatunk be erről a kapcsolatról.

Egy *fizikai modell* az alábbi összetevőkből épül fel: a megfigyelt fizikai jelenségek egy részéből, melyet modellezünk; egy matematikai struktúrából, melyet a fizikai jelenségek matematikai modelljeként értelmezünk; egy oda-vissza megfeleltetésből a matematikai és a fizikai világ között. Az adott matematikai struktúra adja a fizikai modell matematikai részét, a fizikai jelenségek matematikai struktúrába való beágyazása pedig jelenti a fizikai részét.





Nagyon sok fizikai modell létezik, attól függően, hogy mely fizikai jelenséget milyen matematikai struktúrával ír le. Matematikai vizsgálódásunk fő tárgya egy olyan fizikai modellben szerepel, mely bizonyos kvantummechanikai jelenségekre épül. Ezen fizikai modellben szereplő matematikai struktúrát említjük meg, definíció szinten, a fizikai modell többi összetevőjét nem részletezzük.

Az alábbi (egyszerűsített) definíció adja a matematikai keretét a kvantummechanikai matematikai modellek egy részének. A definícióban szereplő matematikai fogalmak közül néhányat még nem definiáltunk, de erre nem is lesz szükség a továbbiakban, ezeket pusztán a teljesség kedvéért említjük meg.

**3.1. Definíció.** Az  $(F, G, L, Q, T_L, T_F, M)$  hetest *absztrakt kvantummechanikai modellnek* nevezzük, ha az alábbiak teljesülnek.

1. Az  $F$  tetszőleges halmaz, melyet a *vizsgálendő események halmazának* nevezünk, és elemeire, mint *eseményekre* utalunk.
2. A  $G$  véges dimenziós Lie-csoport, mely a vizsgált jelenségek szimmetriatulajdonságait fejezi ki.
3. Az  $L$  ortomoduláris háló.
4. A

$$Q : F \rightarrow L \quad E \mapsto Q(E) \quad (3.1)$$

leképezés az eseményekhez egy hálóelemet rendel.

5. A  $T_L$  a  $G$  csoport ábrázolása az  $L$  háló automorfizmusain, azaz

$$T_L : G \rightarrow \text{Aut}(L) \quad g \mapsto T_L(g). \quad (3.2)$$

6. A  $T_F$  a  $G$  csoport ábrázolása az  $F$  halmaz automorfizmusain, azaz

$$T_F : G \rightarrow \text{Aut}(F) \quad g \mapsto T_F(g). \quad (3.3)$$

7. Minden  $g \in G$  és  $E \in F$  esetén

$$T_L(g)(Q(E)) = Q(T_F(g)(E)) \quad (3.4)$$

teljesül, vagyis a

$$\begin{array}{ccc} F & \xrightarrow{Q} & L \\ T_F(g) \downarrow & & \downarrow T_L(g) \\ F & \xrightarrow{Q} & L \end{array} \quad (3.5)$$

diagramm minden  $g \in G$  esetén kommutatív.

8. Az  $M$  az  $L$  hálón értelmezett valószínűségi mértékek egy halmaza.

Az imént definiált absztrakt kvantummechanikai modellnek csak egy részével foglalkozunk. Az események  $F$  halmazát nem vizsgáljuk. Nem célunk egy konkrét mérhető fizikai alkalmazás részletes bemutatása. A megfelelő helyen referenciákra hivatkozunk ezzel kapcsolatban. A jelenségek szimmetriáját leíró  $G$  Lie-csoportot szintén figyelmen kívül hagyjuk. Az  $L$  ortomoduláris háló egy speciális háló lesz esetünkben. A fentiek miatt a  $Q$ ,  $T_L$  és  $T_F$  leképezésekkel sem foglalkozunk. Az  $M$  halmaz lesz az elemzésünk tárgya.

Ezekből látszik, hogy nem célunk a vizsgálandó modellrészlet fizikai relevanciájának és alkalmazásainak az elemzése. A modellrészletünkben az  $L$  és az  $M$  halmaz a következő.

1. Legyen  $L_n$  ( $n \in \mathbb{N}$ ) az  $n$ -dimenziós valós vagy komplex Hilbert-tér ortogonális projekcióinak a halmaza. (Ebben az esetben a Hilbert-tér alterei és ortogonális projekciói kölcsönösen egyértelműen megfeleltethetőek egymásnak.) Jelölje  $P_M$  az  $M$  lineáris altérre való ortogonális projekciót, ekkor  $\text{Ran } P = M$  teljesül. Definiáljuk az alábbi műveleteket az  $L_n$  halmazon.

$$\begin{aligned} Q_1 \wedge Q_2 &= P_{\text{Ran}(Q_1) \cap \text{Ran}(Q_2)} \\ Q_1 \vee Q_2 &= P_{\text{Span}(\text{Ran}(Q_1) \cup \text{Ran}(Q_2))} \\ Q^\perp &= \text{id} - Q, \end{aligned} \quad (3.6)$$

ahol  $\text{Span}$  jelöli a halmaz lineáris burkát és  $\text{id}$  a Hilbert-tér identitás operátorát. Ezekkel a műveletekkel az  $(L, \wedge, \vee, \perp, 0, \text{id})$  hatos moduláris háló. (Ezért ortomoduláris is, azonban  $n > 1$  esetben nem disztributív.) A továbbiakban az egyszerűsítés kedvéért az  $L$  halmaz elemeire mint *eseményekre* hivatkozunk.

Azt mondjuk, hogy a  $Q_1$  és  $Q_2$  *események diszjunktak*, ha  $\text{Ran}(Q_1) \subseteq \text{Ran}(1-Q_2)$  teljesül. A  $Q$  *eseményt elemi eseménynek nevezzük*, ha 1-dimenziós altérre vetít. Az események egy  $(Q_i)_{i \in I}$  rendszerét *teljesnek* nevezzük, ha

$$\bigvee_{i \in I} Q_i = 1 \quad (3.7)$$

teljesül.

2. Azokat a

$$\phi : L \rightarrow [0, 1] \quad P \mapsto \phi(P) \quad (3.8)$$

leképezéseket, melyekre minden  $(P_i)_{i=1, \dots, m}$  páronként diszjunkt véges esemény-rendszerre

$$\phi \left( \bigvee_{i=1}^m P_i \right) = \sum_{i=1}^m \phi(P_i), \quad (3.9)$$

teljesül, valamint  $\phi(1) = 1$ , *valószínűségi mértéknek* nevezzük. Az  $M$  halmaz elemei legyenek az  $L$ -en értelmezett valószínűségi mértékek.

Az általunk vizsgálandó matematikai formalizmust használják például spinnel rendelkező részecskék leírásánál [61, 71, 108, 109].

A klasszikus statisztikai módszerek kvantummechanikai általánosításának a nehézségét alapvetően az okozza, hogy az  $L$  háló nem disztributív. Az  $L$  halmaz a klasszikus esetben szereplő  $X$  alaphalmaz  $\mathcal{B}(X)$   $\sigma$ -algebrájának az általánosítása. Ha ugyanis egy véges sok elemet tartalmazó  $L$  ortomoduláris háló disztributív, akkor  $L$  Boole-algebra, és Stone tétele értelmében ekkor izomorf egy véges halmaz összes részalmazainak a Boole-algebrájával. Vagyis véges  $L$  Boole-háló esetén  $L = \mathcal{P}(X)$  teljesül és ebben az esetben  $M$  bizonyos diszkrét eloszlások halmaza. Ekkor az (1.1.) példában bemutatott diszkrét eloszlás esetét kapjuk vissza. (Ha  $L$  végtelen Boole-algebra, akkor egy halmaztesttel izomorf.) Tehát az általánosítás fő nehézsége az, hogy nem disztributív  $L$  esetén nincs  $X$  „alaptér”.

Az  $M$  halmaz szerkezete jól jellemezhető az alábbi Gleason-tól származó tétellel.

**3.1. Tétel.** (Gleason tétele:) Legyen  $\mathcal{H}$  véges dimenziós valós vagy komplex Hilbert-tér, melyre  $\dim \mathcal{H} = n \neq 2$  teljesül, jelölje  $L_n$  a Hilbert-tér ortogonális projektorainak

a hálóját és legyen  $\phi : L_n \rightarrow [0, 1]$  valószínűségi mérték. Ekkor létezik egyetlen egy  $D$   $n \times n$ -es pozitív, egységnyomú önadjungált operátor, melyre

$$\phi(P) = \text{Tr} DP \quad \forall P \in L_n \quad (3.10)$$

teljesül. Továbbá minden  $D$  pozitív, egységnyomú önadjungált  $n \times n$ -es operátor esetén a (3.10) képlet valószínűségi mértéket definiál az  $L_n$  hálón.

Gleason eredeti tétele szeparábilis Hilbert-térre vonatkozik. A tétel általánosításai megtalálhatóak a [27] könyvben. (Érdekesség, hogy a tétel könnyen visszavezethető a három dimenziós esetre, ami azonban messze nem triviális.)

Ez a tétel adja az alapját a következő részben definiálandó kvantummechanikai állapottérnek.

### 3.2. A kvantummechanikai formalizmus

Az alábbi definícióval általánosítjuk a  $\mathcal{P}_n$  halmazt.

**3.2. Definíció.** Az  $n$  dimenziós valós (illetve komplex-) Hilbert-tér önadjungált, pozitív, egységnyomú operátorainak a halmazát  *$n$ -dimenziós valós (illetve komplex) állapottérnek* nevezzük, melynek az elemei a *valós (illetve komplex) állapotok*. Az állapotot gyakran *sűrűségi mátrixnak* is nevezik. A valós illetve komplex  $n$ -dimenziós állapottér belsejét egyformán  $\mathcal{M}_n^+$  jelöli, amennyiben ez nem okoz félreértést.

Fontos megemlíteni, hogy a két dimenziós Hilbert-tér esetén az  $M$  halmaznak nem minden eleme reprezentálható állapottal.

Az eddigiek értelmében annak a valószínűsége, hogy egy  $D \in M$  állapotban lévő kvantummechanikai rendszerben a  $P \in L$  esemény bekövetkezik  $\text{Tr} DP$ . Definícióból következik, hogy az állapottér konvex, zárt halmaz, ez adja a lehetőséget a további osztályozásukra.

**3.3. Definíció.** Az állapottér extrémális pontjainak a halmazát *tiszta állapotoknak* nevezzük, míg nem extrémális pontjait *kevert állapotoknak*.

Ezek szerint az  $\mathcal{M}_n^+$  halmaznak egyetlen tiszta állapot sem az eleme. Igazolható, hogy egy  $D$  állapot rangja ( $= \dim \text{Ran } D$ ), éppen azt adja meg, hogy a  $D$  állapot legalább hány tiszta állapot konvex kombinációjaként állítható elő.

A mérhető fizikai mennyiségek fogalmát az alábbi gondolatmenettel terjeszthetjük ki a kvantumos esetre. Egy fizikai mennyiségre, mint például az energiára, úgy gondolunk, hogy ahhoz létezik az eseményeknek egy páronként diszjunkt  $(P_i)_{i \in I}$  teljes

rendszere, melyben a fizikai mennyiség „tisztán” mérhető, és egy-egy ilyen állapotban a  $(\lambda_i)_{i \in I}$  valós értékeket veszi fel. A mennyiséghez tartozó ezen információkat egy

$$A = \sum_{i \in I} \lambda_i P_i \quad (3.11)$$

önadjungált operátorral lehet kódolni. Továbbá minden  $A$  önadjungált operátor lényegében egyértelműen felírható a (3.11) alakban. Ez motiválja a következő definíciót.

**3.4. Definíció.** Az  $n$ -dimenziós valós vagy komplex Hilbert-térrel modellezett kvantummechanikai rendszerek esetén a valós illetve komplex önadjungált mátrixok halmazát *mérhető fizikai mennyiségeknek* (vagy röviden *fizikai mennyiségeknek* illetve *obszervábiliseknek*) nevezzük. Ha a rendszer a  $D$  sűrűségi mátrixszal jellemezhető, akkor ebben az állapotban a  $H$  *fizikai mennyiség várható értéke*

$$E(H) = \text{Tr } DH, \quad (3.12)$$

*k-adik momentuma*

$$E_k(H) = \text{Tr } DH^k \quad (3.13)$$

és *szórása*

$$\sigma(H) = \sqrt{\text{Tr } DH^2 - (\text{Tr } DH)^2}. \quad (3.14)$$

Az eddig bevezetett definíciókról könnyen látható, hogy a klasszikus diszkrét eset általánosításai. Tekintsük ugyanis az  $n$ -dimenziós Hilbert-teret. Ekkor, ha a diagonális mátrixokra szorítkozunk, a projekciók halmaza azonosítható egy  $n$  pontból álló halmaz részhalmazával, egy állapotot leíró sűrűségi mátrix pedig egy sűrűségfüggvénynek felel meg, valamint egy fizikai mennyiség valójában egy függvény az  $n$  elemű halmazon. A fenti definícióban szereplő képletek is visszaadják a klasszikus esetet ekkor.

Vizsgáljuk meg, hogy alacsony  $n$ -dimenziós Hilbert-terek esetén hogyan változik az  $\mathcal{M}_n^+$  halmaz szerkezete. Az  $n = 0, 1$  eset triviális.

**3.1. Példa.** Az  $n = 2$  esetben a komplex állapotokra az alábbi, *Stokes-paraméterezést* fogjuk használni. Minden  $D \in \mathcal{M}_n^+$  állapot, felírható

$$D = \frac{1}{2}(x\sigma_1 + y\sigma_2 + z\sigma_3 + I) \quad (3.15)$$

alakban, ahol

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \sigma_2 = \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix} \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad (3.16)$$

a *Pauli-mátrixok*, valamint  $I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$  az egységmátrix és  $x, y, z \in \mathbb{R}$ . Igazolható, hogy a (3.15) egyenlet pontosan akkor határoz meg állapotot, ha

$$x^2 + y^2 + z^2 \leq 1 \quad (3.17)$$

teljesül. Ezek alapján az  $\mathcal{M}_n^+$  tér az  $n = 2$  esetben azonosítható a háromdimenziós euklideszi tér nyílt egységgömbjével.  $\square$

Magasabb dimenzióban az állapotok tere, meglehetősen bonyolult alakzat lesz.

### 3.3. Állapotok entrópiája és majorizációja

Adott  $D \in \mathcal{M}_n^+$  állapot entrópiáját hasonló módon lehet értelmezni, mint a klasszikus esetben (lásd az 1.8. definíciót). Az entrópia ilyen irányú kiterjesztését Neumann javasolta először 1927-ben [76].

**3.5. Definíció.** A  $D \in \mathcal{M}_n^+$  állapot *Neumann-féle entrópiája* a

$$S(D) = -\text{Tr } D \log D \quad (3.18)$$

mennyiség.

Ez a mennyiség nem más, mint a  $D$  sajátértékeiből álló diszkrét eloszlás Shannon-féle entrópiája.

A Neumann-féle entrópia tulajdonságai részletes bizonyításokkal együtt megtalálhatók Petz és Hiai [51], valamint Petz és Ohya [77] könyvében vagy a [78] cikkben.

Az 1.2. részben láttuk, hogy fontos szerepet töltenek be a maximális entrópiájú állapotok. Az entrópia maximalizációjának a problémáját megoldottuk a klasszikus esetben (az 1.4. példában), azon feltétel mellett, hogy rendszer energiájának a várható értéke adott. Ezen megoldásokat neveztük Gibbs-állapotoknak. A kvantummechanikai esetben Neumann oldotta meg a fenti optimalizációs problémát. A megoldás menetét mellőzve, csak a végeredményt említjük.

**3.2. Tétel.** *Tekintsük az  $S : \mathcal{M}_n^+ \rightarrow \mathbb{R}$  entrópiát, és legyen  $H$  tetszőleges  $n \times n$ -es önadjungált operátor. A  $\text{Tr } DH = e$  feltétel mellett az entrópia akkor maximális, ha*

$$D = \frac{e^{-\beta H}}{\text{Tr } e^{-\beta H}} \quad (3.19)$$

*teljesül, valamilyen  $\beta \in \mathbb{R}$  mellett. A fenti feltétel egyértelműen meghatározza a  $\beta$  paramétert.*

A klasszikus esethez hasonlóan a maximális entrópiájú állapotok adják a Gibbs-állapotokat.

**3.6. Definíció.** Adott  $\beta$  valós paraméter és  $H$  önadjungált mátrix esetén

$$R_{(H)}^{(\beta)} = \frac{e^{-\beta H}}{\text{Tr } e^{-\beta H}} \quad (3.20)$$

jelöli a *Gibbs-állapotot*.

A definícióban szereplő  $\beta$  paraméter az *inverz hőmérséklet* a  $\beta = \frac{1}{kT}$  összefüggés miatt.

Az alábbi tételből kiderül, hogy az entrópia konkáv függvény.

**3.3. Tétel.** Minden  $D_1, D_2 \in \mathcal{M}_n^+$  állapot és  $\lambda \in [0, 1]$  paraméter esetén

$$\lambda S(D_1) + (1 - \lambda)S(D_2) \leq S(\lambda D_1 + (1 - \lambda)D_2) \quad (3.21)$$

teljesül.

A majorizációs reláció is kiterjeszthető az állapottérre, a klasszikus esetre vonatkozó 1.11. definíció alapján.

**3.7. Definíció.** Legyen  $D_1, D_2 \in \mathcal{M}_n^+$  állapot, valamint jelölje  $(\mu_1^\downarrow, \dots, \mu_n^\downarrow)$  és  $(\lambda_1^\downarrow, \dots, \lambda_n^\downarrow)$  a  $D_1$  és  $D_2$  állapot sajátértékeit, csökkenő sorrendbe rendezve. Azt mondjuk, hogy  $D_1$  *majorálja*  $D_2$ -t (vagy  $D_2$  *kevertebb*  $D_1$ -nél), ha minden  $k = 1, \dots, n$  esetén

$$\sum_{i=1}^k \lambda_i^\downarrow \leq \sum_{i=1}^k \mu_i^\downarrow \quad (3.22)$$

teljesül. Ezen reláció fennállását  $D_2 \prec D_1$  jelöli.

Az azonos sajátértékkel rendelkező állapotot minden állapot majorálja. Ezért ezt az állapotot a *legkevertebb állapotnak* nevezzük.

A klasszikus esetre vonatkozó 1.9. tétel, mely szerint a magasabb hőmérséklethez tartozó Gibbs-állapot kevertebb, szintén érvényes.

**3.4. Tétel.** *Wehrl tétele [111]:* Adott  $H$  önadjungált operátor és  $\beta_1 < \beta_2$  paraméterek esetén

$$R_{(H)}^{(\beta_1)} \prec R_{(H)}^{(\beta_2)} \quad (3.23)$$

teljesül.

Az 1.10. tétel kissé módosított formában szintén érvényes [14, 81].

**3.5. Tétel.** Legyen  $D_1, D_2 \in \mathcal{M}_n^+$ . Ekkor az alábbiak ekvivalensek.

1. A  $D_1 \prec D_2$  reláció teljesül.
2. Léteznek olyan  $P_0, \dots, P_k \in \mathcal{M}_n^+$  állapotok, melyekre  $P_0 = D_1, P_k = D_2$ , továbbá

$$P_0 \prec \dots \prec P_k \quad (3.24)$$

teljesül, valamint minden  $i = 0, \dots, k-1$  esetén a  $P_i$  állapot sajátértékei a  $P_{i+1}$  állapot sajátértékeinek  $T$ -transzformációjával kapható meg.

3. Léteznek  $P_0, \dots, P_m \in \mathcal{M}_n^+$  állapotok, melyekre  $P_0 = D_1, P_m = D_2$ , továbbá

$$P_0 \prec \dots \prec P_m \quad (3.25)$$

teljesül, valamint minden  $i = 0, \dots, m-1$  esetén létezik  $H_i$  önadjungált operátor,  $\beta_{1,i}$  és  $\beta_{2,i}$  paraméterek, hogy  $P_i = R_{(H_i)}^{(\beta_{1,i})}$  és  $P_{i+1} = R_{(H_i)}^{(\beta_{2,i})}$  teljesül.

A klasszikus esethez hasonlóan itt is nagyobb a rendezetlensége a kevertebb állapotoknak, azaz  $D_1 \prec D_2$  esetén  $S(D_2) \leq S(D_1)$  teljesül.

### 3.4. A Fisher-információ általánosítása

Statisztikai modellek esetére definiáltuk a Fisher-féle információt. Ezért először a statisztikai modell fogalmát általánosítjuk.

**3.8. Definíció.** A  $\mathcal{Q} = (L_n, \mathcal{Q}, \Xi)$  hármast  $m$ -dimenziós kvantummechanikai statisztikai modellnek nevezzük, ha

1.  $L_n$  az  $n$ -dimenziós valós vagy komplex Hilbert-tér projektorainak a halmaza,
2.  $\mathcal{Q} \subseteq \mathcal{M}_n^+$  és  $\Xi \subseteq \mathbb{R}^m$  teljesül, továbbá  $\Xi$  összefüggő nyílt halmaz,
3. létezik egy

$$i : \Xi \rightarrow \mathcal{Q} \quad \vartheta \mapsto D_\vartheta$$

$C^\infty$ -beli bijekció.

Általános kvantummechanikai statisztikai modell esetén a klasszikus Fisher-féle információ tulajdonságai alapján próbáljuk a kvantumos megfelelőjét megtalálni. Ehhez az 1.1. részben említett tételeket vesszük sorra.

Az 1.2. tételt (mely a Fisher-féle információ monotonitását mondja ki) nehéz átfogalmazni a kvantumos esetre, a benne szereplő  $f : X \rightarrow Y$  elégséges statisztika miatt, ugyanis kvantumos esetben nincs  $X, Y$  alaptér.



Az 1.3. tételnél, mely hasonlót fogalmaz meg mint az 1.2 tétel, annyival jobb a helyzet, hogy a benne szereplő átmenetvalószínűség fogalmát lehet általánosítani a kvantumos esetre, azonban még ekkor is meglehetősen sok olyan függvény létezik, mely monoton az átmenetvalószínűségekre nézve.

Az 1.4. és 1.5. tételek (additivitás, Cramer-Rao egyenlőtlenség) a klasszikus Fisher-féle információ hasznos tulajdonságai, azonban ezek alapján sem lehet jó definíciót adni a kvantumos esetre.

Az 1.8. tétel (mely szerint a legtöbb relatív entrópia a Fisher-féle információt generálja) hasznos tulajdonság, azonban a relatív entrópiát még nem értelmeztük kvantumos esetre, valamint nem igaz, hogy minden relatív entrópia a Fisher-féle információt generálja.

A 2.17. tétel (mely szerint a Fisher-féle információ pozitív számszorzó erejéig egyértelmű) általánosítható a kvantumos esetre. A tételben ugyan szerepel az alapterek közötti  $f : X_n \rightarrow X_m$  szürjektív leképezés fogalma, de igazából csak az általa indukált  $\tilde{f} : \mathcal{P}_n \rightarrow \mathcal{P}_m$  leképezés kap kulcsszerepet.

A Fisher-féle információ ilyen irányú kiterjesztésében Cencov és Morozova tette meg az első lépéseket. 1996-ban Petznek sikerült a kvantumos Fisher-információt ilyen módon jellemeznie [82]. Ezen eredmények áttekintéséhez azonban további definíciókat, valamint az  $\mathcal{M}_n^+$  differenciálható struktúráját kell bevezetnünk.

**3.9. Definíció.** Jelölje  $M_n$  az  $n \times n$ -es komplex vagy valós mátrixok terét és jelölje  $M_m(M_n)$  azon  $m \times m$ -es mátrixokat, melyek elemei  $n \times n$ -es mátrixok. A  $T : M_n \rightarrow M_m$  lineáris leképezésről azt mondjuk, hogy *pozitív*, ha pozitív operátorhoz pozitív operátort rendel.

A  $T : M_n \rightarrow M_m$  lineáris leképezésről azt mondjuk, hogy *teljesen pozitív*, ha minden  $k \in \mathbb{N}$  esetén a

$$T^{(k)} : M_k(M_n) \rightarrow M_k(M_m) \quad [A_{ij}] \mapsto T^{(k)}([A_{ij}]) = [T(A_{ij})] \quad (3.26)$$

lineáris leképezés pozitív.

A  $T : M_n \rightarrow M_m$  lineáris leképezésről azt mondjuk, hogy *sztochasztikus leképezés*, ha megőrzi a mátrix nyomát és teljesen pozitív.

A teljesen pozitív leképezések főbb tulajdonságai és reprezentációs tételei megtalálhatók [8, 20, 57, 68, 97] munkákban. Ezek közül csak egy egyszerűsített tételt, a Kraus-tételt emeljük ki.

**3.6. Tétel.**  $A T : M_n \rightarrow M_m$  leképezés pontosan akkor teljesen pozitív, ha létezik véges sok  $V_i : M_n \rightarrow M_m$  operátor úgy, hogy

$$T(A) = \sum_{i=1}^N V_i A V_i^* \quad \forall A \in M_n \quad (3.27)$$

teljesül. A teljesen pozitív  $T$  leképezés pontosan akkor nyomtartó, ha  $\sum_{i=1}^N V_i V_i^* = id$ .

A sztochasztikus leképezésekről bővebben Alberti és Uhlmann [1] monográfiájában vagy Kraus [57] könyvében lehet olvasni.

Minden  $n \in \mathbb{N}$  esetén a  $\mathcal{M}_n^+$  halmaz az  $\mathbb{R}$  tér megfelelő hatványhalmazában egy összefüggő nyílt részhalmazzal azonosítható. Ennek segítségével  $\mathcal{M}_n^+$  differenciálható sokasággá tehető.

A koordinátázás megadásához definiáljuk az alábbi mátrixokat

$$\sigma_r^{(k,l)} = \frac{1}{2} F_{kl}, \quad \sigma_c^{(k,l)} = \frac{1}{2} H_{kl} \quad 1 \leq k < l \leq n \quad (3.28)$$

$$\sigma_d^{(k)} = E_{kk} \quad 1 \leq k \leq n, \quad (3.29)$$

ahol  $H_{kl} = -i E_{kl} + i E_{lk}$  a komplex mátrixegység. Komplex állapotok esetében a

$$\phi : \mathcal{M}_n^+ \rightarrow \mathbb{R}^{n^2-1} \quad D \mapsto \phi(D) \quad (3.30)$$

$$\begin{aligned} \phi(D) = & (\text{Tr } D\sigma_d^{(1)}, \dots, \text{Tr } D\sigma_d^{(n-1)}, \text{Tr } D\sigma_r^{(1,2)}, \dots, \text{Tr } D\sigma_r^{(1,n)}, \dots \\ & \dots, \text{Tr } D\sigma_r^{(k,k+1)}, \dots, \text{Tr } D\sigma_r^{(k,n)}, \dots, \text{Tr } D\sigma_r^{(n,n)}, \dots, \text{Tr } D\sigma_c^{(1,2)}, \\ & \dots, \text{Tr } D\sigma_c^{(1,n)}, \dots, \text{Tr } D\sigma_c^{(k,k+1)}, \dots, \text{Tr } D\sigma_c^{(k,n)}, \dots, \text{Tr } D\sigma_c^{(n,n)}) \end{aligned}$$

koordinátázást, valamint valós állapotok esetében a

$$\phi : \mathcal{M}_n^+ \rightarrow \mathbb{R}^{\frac{(n+2)(n-1)}{2}} \quad D \mapsto \phi(D) \quad (3.31)$$

$$\begin{aligned} \phi(D) = & (\text{Tr } D\sigma_d^{(1)}, \dots, \text{Tr } D\sigma_d^{(n-1)}, \text{Tr } D\sigma_r^{(1,2)}, \dots, \text{Tr } D\sigma_r^{(1,n)}, \dots \\ & \dots, \text{Tr } D\sigma_r^{(k,k+1)}, \dots, \text{Tr } D\sigma_r^{(k,n)}, \dots, \text{Tr } D\sigma_r^{(n,n)}) \end{aligned}$$

leképezést *kanonikus koordinátázásnak* (vagy *kanonikus paraméterezésnek*) nevezzük. Ezek szerint az  $\mathcal{M}_n^+$  halmaz egy koordinátázással lefedhető. (Amennyiben diagonális állapotokra szorítkozunk, visszakapjuk a diszkrét állapotok 2.1. pontban említett koordinátázását.)

A kanonikus koordinátázás annyira természetes, hogy a továbbiakban az  $\mathcal{M}_n^+$  térre gyakran úgy hivatkozunk, mint a megfelelő dimenziós euklideszi tér részhalmazára.

Adott  $D_0 \in \mathcal{M}_n^+$  pontban az érintőteret azonosíthatjuk az önadjungált, nulla nyomú,  $n \times n$ -es mátrixok  $\mathcal{M}_n$  terével. [Valós (komplex) állapotok esetén az  $\mathcal{M}_n$

érintőtérben valós (komplex) mátrixok szerepelnek.] Adott  $f : \mathcal{M}_n^+ \rightarrow \mathbb{R}$  sima függvény,  $D_0 \in \mathcal{M}_n^+$  állapot és  $X \in \mathcal{M}_n$  érintővektor esetén a

$$(Xf)(D_0) = \left. \frac{d f(D_0 + tX)}{dt} \right|_{t=0} \quad (3.32)$$

kifejezéssel értelmezzük a deriváció hatását az  $f$  függvényen (összhangban a (2.8) képlettel). Jelölje  $T\mathcal{M}_n$  az  $\mathcal{M}_n^+$  sokaság érintőnyalábját.

Az  $\mathcal{M}_n^+$  sokaságot többféle metrikával tehetjük Riemann-sokasággá. Például a

$$K : \mathcal{M}_n^+ \rightarrow \text{Lin}(T\mathcal{M}_n \times T\mathcal{M}_n, \mathbb{R}) \quad (D, X, Y) \mapsto \text{Tr } DXY \quad (3.33)$$

leképezés Riemann-metrikát definiál.

Az állapot tér, mint Riemann-sokaság, új eszközt jelent az állapotok tulajdonságainak a vizsgálatában. Az egyik fontos feladat megtalálni, hogy mely Riemann-metrikák lehetnek relevánsak fizikai szempontokból, a másik, az adott Riemann-metrika esetén a differenciálgeometriai fogalmak (például görbület, konnexió) megfelelő fizikai interpretálása. A második célkitűzés túlmutat a matematika keretein, azonban adott Riemann-geometriák részletes matematikai elemzése segítséget jelent ebben a kérdésben. Az állapot téren értelmezett Riemann-metrikák statisztikus fizikai alkalmazhatóságát Balian [9] és Streater [98] vizsgálták, az átlagtér elméletben pedig Tanaka [101] alkalmazta sikerrel.

**3.10. Definíció.** Az  $(\mathcal{M}_n^+, K^{(n)})_{n \in \mathbb{N}}$  Riemann-sokaságok egy családját, melyre minden  $n, m \in \mathbb{N}$  paraméterekre, minden  $T : M_n \rightarrow M_m$  sztochasztikus leképezésre, minden  $D \in \mathcal{M}_n^+$  állapotra és minden  $X \in \mathcal{M}_n$  érintőtérbeli vektorra

$$K_{T(D)}^{(m)}(T(X), T(X)) \leq K_D^{(n)}(X, X) \quad (3.34)$$

teljesül, *monoton metrikák családjának* nevezzük.

A definícióban szereplő „monoton” jelző félreérthető lehet olykor (nem fejezi ki hűen a definíciót), azonban csak ebben az értelemben vett monoton metrikák családjai fognak szerepelni a későbbiekben.

**3.11. Definíció.** Egy  $f : \mathbb{R} \rightarrow \mathbb{R}$  függvény *operátormonoton*, ha minden  $n \in \mathbb{N}$  számra és  $A, B \in M_n$  önadjungált mátrixra  $A \leq B$  esetén  $f(A) \leq f(B)$  teljesül.

Jelölje  $\text{Lin}(M_n)$  az  $A : M_n \rightarrow M_n$  lineáris leképezések halmazát. A  $\text{Lin}(M_n)$  téren értelmezzük a

$$\langle \cdot, \cdot \rangle : \text{Lin}(M_n) \times \text{Lin}(M_n) \rightarrow \mathbb{C} \quad (A, B) \mapsto \text{Tr } A^*B \quad (3.35)$$

*Hilbert–Schmidt-féle (vagy kanonikus) skalárszorozást.* Minden  $D \in M_n$  mátrixhoz definiáljuk az  $L_{n,D}, R_{n,D} \in \text{Lin}(M_n)$  leképezéseket a

$$L_{n,D}(A) = DA \quad R_{n,D}(A) = AD \quad (3.36)$$

képletekkel, melyeket a *balról* illetve *jobbról való szorzás operátorának* nevezünk. Ha  $D \in \mathcal{M}_n^+$  állapot, akkor az  $L_{n,D}$  és  $R_{n,D}$  leképezés önadjungált, az

$$\begin{aligned} \langle L_{n,D}A, B \rangle &= \langle DA, B \rangle = \text{Tr}(DA)^*B = \text{Tr} A^*D^*B = \\ &= \text{Tr} A^*DB = \langle A, DB \rangle = \langle A, L_{n,D}B \rangle \\ \langle R_{n,D}A, B \rangle &= \langle AD, B \rangle = \text{Tr}(AD)^*B = \text{Tr} D^*A^*B = \\ &= \text{Tr} A^*BD = \langle A, BD \rangle = \langle A, R_{n,D}B \rangle \end{aligned} \quad (3.37)$$

egyenletek alapján. Sőt igazolható, hogy az  $L_{n,D}$  és  $R_{n,D}$  leképezések sajátértékei megegyeznek a  $D$  sajátértékeivel, azonban a  $D$  minden  $k$ -szoros sajátértéke,  $nk$ -szoros sajátértéke lesz az  $L_{n,D}, R_{n,D}$  operátoroknak.

Ezen definíciók segítségével lehet megfogalmazni Cencov és Morozova Fisher-információ kvantumossal általánosítására vonatkozó 1990-es eredményét, valamint a teljes jellemzést adó Petz-tételt.

**3.7. Tétel.** (*Cencov–Morozova [74]:*) Legyen  $(\mathcal{M}_n^+)_{n \in \mathbb{N}, K^{(n)}}$  monoton metrikák családja. Ekkor létezik olyan  $c : ]0, \infty]^2 \rightarrow ]0, \infty]$  folytonos függvény és pozitív  $C$  állandó, hogy minden  $D \in \mathcal{M}_n^+$  állapotra  $D = \sum_{i=1}^n \lambda_i E_{ii}$  esetén

$$K_D^{(n)}(A, A) = C \sum_{i=1}^n \frac{1}{\lambda_i} A_{ii}^2 + \sum_{1 \leq i < j \leq n} |A_{ij}|^2 c(\lambda_i, \lambda_j) \quad \forall A \in T_D \mathcal{M}_n^+ . \quad (3.38)$$

Továbbá minden pozitív  $\mu, \lambda$  és  $t$  paraméterre  $c(\lambda, \mu) = c(\mu, \lambda)$ ,  $c(\lambda, \lambda) = C\lambda^{-1}$  és  $c(t\lambda, t\mu) = t^{-1}c(\lambda, \mu)$  teljesülnek.

A fenti eredmény hiányossága, hogy nem mondja meg, mely  $c$  függvény esetén lehet Riemann-metrikát származtatni a (3.38) képlet segítségével. Ezt a hiányosságot pótolja Petz alábbi tétele, mely szerint bizonyos operátormonoton függvények segítségével lehet monoton metrikákat megadni.

**3.8. Tétel.** (*Petz osztályozási tétele [82]:*) A monoton metrikák családja és az olyan operátormonoton  $f : \mathbb{R}^+ \rightarrow \mathbb{R}$  függvények között, melyekre minden pozitív  $x$  esetén  $f(x) = xf(x^{-1})$  teljesül, létezik kölcsönösen egyértelmű megfeleltetés. Minden ilyen  $f$  függvényre a

$$K^{(n),f} : \mathcal{M}_n^+ \times \mathcal{M}_n \times \mathcal{M}_n \rightarrow \mathbb{R} \quad (D, X, Y) \mapsto K_D^{(n),f}(X, Y) \quad (3.39)$$

$$K_D^{(n),f}(X, Y) = \text{Tr} \left( X \left( R_{n,D}^{\frac{1}{2}} f(L_{n,D} R_{n,D}^{-1}) R_{n,D}^{\frac{1}{2}} \right)^{-1} (Y) \right) \quad (3.40)$$

kifejezés a  $(K^{(n),f}, \mathcal{M}_n^+)_{n \in \mathbb{N}}$  monoton metrikák egy családját generálja. Továbbá a monoton metrikák minden családjára előáll a (3.39) formában, valamilyen alkalmas  $f$  függvénnyel.

**3.12. Definíció.** Az  $(\mathcal{M}_n^+, K^{(n)})$  Riemann-sokaságban szereplő  $K^{(n)}$  metrikát *monoton metrikának* nevezzük, ha van olyan operátormonoton  $f : \mathbb{R}^+ \rightarrow \mathbb{R}$  függvény, amellyel pozitív  $x$  számokra  $f(x) = xf(x^{-1})$  teljesül, és amely a  $K^{(n)}$  metrikát indukálja a (3.39) képlet segítségével.

A  $K^{(n)}$  metrikát még az  $(L_n, \mathcal{M}_n^+)$  kvantummechanikai statisztikai modellhez tartozó *kvantummechanikai Fisher-féle információ*nak is nevezzük.

### 3.5. Kvantummechanikai Fisher-féle információ

Petz tétele értelmében a monoton metrikák bizonyos operátormonoton függvények segítségével indexelhetők. Az operátormonoton függvényeknek az 1930-as években Löwner adta meg a jellemzését [69], a [6, 14, 24, 42] művekben modernebb megközelítésmóddal lett újra feldolgozva és kibővítve a témakör. Most Löwner klasszikus eredményeit, illetve egyszerű, de a jelen esetben nagyon hasznos következményeit említjük.

**3.13. Definíció.** Legyen  $f : [0, \infty] \rightarrow \mathbb{R}$  operátormonoton függvény. Az  $f^\setminus(x) = xf(x^{-1})$  függvényt az  $f$  függvény *transzponáltjának*, és az  $f^\perp(x) = \frac{x}{f(x)}$  függvényt az  $f$  függvény *duálisának* nevezzük. Azt mondjuk, hogy az  $f$  függvény *szimmetrikus*, ha  $f = f^\setminus$  teljesül, illetve az  $f$  függvény *normalizált*, ha  $f(1) = 1$  teljesül.

A transzponálás és a duális képzés involutív művelet, vagyis  $f = f^{\perp\perp}$  és  $f = f^{\setminus\setminus}$ .

Adott  $I \subseteq \mathbb{R}$  intervallum esetén az operátormonoton  $f : I \rightarrow \mathbb{R}$  függvények halmazát jelölje  $\mathcal{F}_I$ , a normalizált  $\mathcal{F}_I$ -beli elemeket ( $1 \in I$  esetén) jelölje  $\mathcal{F}_I^{(n)}$ . Az  $I$  halmazon értelmezett pozitív Radon-mértékek halmazát jelölje  $\mathcal{G}_I$ . Egy  $\mu \in \mathcal{G}_I$  mértékről azt mondjuk, hogy *normalizált*, ha  $\mu(I) = 1$  teljesül. A normalizált  $\mathcal{G}_I$ -beli mértékek halmazát jelölje  $\mathcal{G}_I^{(n)}$ .

Hasegawa igazolta, hogy a szimmetrikus operátormonoton függvények köréből, nem vezet ki a duálisképzés [44].

**3.9. Tétel.** Az  $f : [0, \infty] \rightarrow \mathbb{R}$  szimmetrikus, operátormonoton függvény duálisa is operátormonoton.

**3.10. Tétel.** (Löwner tétele:[58]) A

$$\begin{aligned} \phi : \mathcal{G}_{[0,\infty]} &\rightarrow \mathcal{F}_{[0,\infty]} & \mu &\mapsto f_\mu \\ f_\mu(x) &= \int_0^\infty \frac{x(1+t)}{x+t} d\mu(t) \end{aligned} \tag{3.41}$$

leképezés kölcsönösen egyértelmű megfeleltetés a  $\mathcal{G}_{[0,\infty]}$  és a  $\mathcal{F}_{[0,\infty]}$  halmaz között. A  $\phi$  leképezés megszorítása az  $\mathcal{G}_{[0,\infty]}^{(n)}$  halmazra kölcsönösen egyértelmű megfeleltetést létesít a  $\mathcal{G}_{[0,\infty]}^{(n)}$  és az  $\mathcal{F}_{[0,\infty]}^{(n)}$  halmaz között.

Az alábbi reprezentációs tétel megtalálható Giblisco és Isola [39] cikkében.

### 3.11. Tétel. A

$$\phi : \mathcal{G}_{[0,1]} \rightarrow \mathcal{F}_{[0,\infty]} \quad \mu \mapsto f_\mu \quad (3.42)$$

$$f_\mu(x) = \int_0^1 \frac{x}{(1-t)x+t} d\mu(t) \quad (3.43)$$

leképezés kölcsönösen egyértelmű megfeleltetés a  $\mathcal{G}_{[0,1]}$  és a  $\mathcal{F}_{[0,\infty]}$  halmazok között. A  $\phi$  leképezés megszorítása a  $\mathcal{G}_{[0,1]}^{(n)}$  halmazra kölcsönösen egyértelmű megfeleltetést létesít a  $\mathcal{G}_{[0,1]}^{(n)}$  és az  $\mathcal{F}_{[0,\infty]}^{(n)}$  halmaz között. Továbbá az  $f_\mu$  függvény pontosan akkor szimmetrikus, ha  $\mu([0, s]) = \mu([1-s, 1])$  teljesül minden  $0 \leq s \leq 1$  esetén.

A  $\mathcal{G}_{[a,b]}$  halmaz azon elemeit, melyekre  $\mu([a, s]) = \mu([b-s, b])$  teljesül minden  $a \leq s \leq b$  esetén, szimmetrikus mértékeknek nevezzük. A szimmetrikus mértékek halmazát jelölje  $\mathcal{G}_{[a,b]}^{(S)}$ . A szimmetrikus normalizált mértékek halmaza legyen  $\mathcal{G}_{[a,b]}^{(S,n)}$ .

**3.14. Definíció.** A  $c : [0, \infty]^2 \rightarrow \mathbb{R}$  függvényt Cencov–Morozova-féle függvénynek nevezzük, ha létezik olyan  $f \in \mathcal{F}_{[0,\infty]}$  függvény, melyre minden pozitív  $x, y$  szám esetén

$$c(x, y) = \frac{1}{yf\left(\frac{x}{y}\right)} \quad (3.44)$$

teljesül.

Minden  $f \in \mathcal{F}_{[0,\infty]}$  függvényhez egy  $c_f$  Cencov–Morozova-féle függvényt lehet asszociálni a definícióban szereplő (3.44) képlettel. Az elnevezést az indokolja, hogy minden ilyen  $c$  függvény monoton metrikát generál a (3.38) képlettel.

Az  $f \in \mathcal{F}_{[0,\infty]}$  függvény pontosan akkor szimmetrikus, ha a  $c_f$  függvény is szimmetrikus (azaz minden pozitív  $x, y$  számra  $c_f(x, y) = c_f(y, x)$  teljesül). Az  $\mathcal{F}_I$  és  $\mathcal{F}_I^{(n)}$  függvényhalmaz szimmetrikus elemeinek a halmazát jelölje  $\mathcal{F}_I^{(S)}$  és  $\mathcal{F}_I^{(S,n)}$ .

**3.12. Tétel.** ([14]:) Minden  $\mu \in \mathcal{G}_{[0,1]}$  mértékre a

$$f^{(\mu)}(x) = \int_0^1 x^t d\mu(t) \quad (3.45)$$

képlettel értelmezett függvény  $\mathcal{F}_{[0,\infty]}$ -beli, azonban nem minden  $\mathcal{F}_{[0,\infty]}$ -beli függvény állítható elő ilyen alakban. Továbbá a  $\mu \in \mathcal{G}_{[0,1]}^{(S)}$  és a  $\nu \in \mathcal{G}_{[0,1]}^{(n)}$  mértékre  $f^{(\mu)} \in \mathcal{F}_{[0,\infty]}^{(S)}$  és  $f^{(\nu)} \in \mathcal{F}_{[0,\infty]}^{(n)}$  teljesül.

A fenti reprezentációs tételből adódik az alábbi tétel.

**3.13. Tétel.** Minden  $\mu \in \mathcal{G}_{[0,1]}$  mérték esetén a

$$c^{(\mu)}(x, y) = \int_0^1 \frac{1}{x^{1-t}y^t} d\mu(t) \quad (3.46)$$

képlettel értelmezett függvény Cencov–Morozova-féle, azonban nem minden Cencov–Morozova-féle függvény állítható elő ilyen alakban.

A Cencov–Morozova-féle függvényekkel is ki lehet fejezni Petz-tételét. Ennek a megértéséhez azonban egy új operátorkalkulust kell alkalmazni.

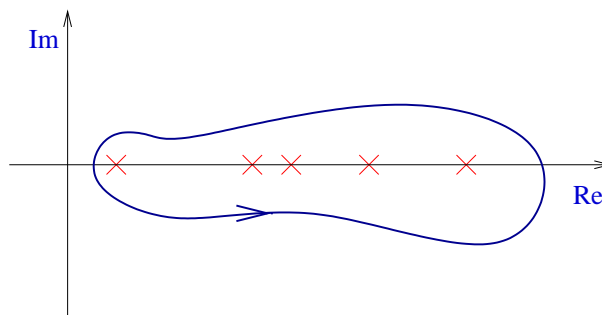
Komplex függvénytanból ismert, hogy amennyiben egy  $f : \mathbb{C} \rightarrow \mathbb{C}$  komplex analitikus függvény az  $\mathbb{R}^+$  halmaz egy  $\mathcal{D}$  környezetén van értelmezve, akkor minden  $\rho \in \mathbb{R}$  esetén

$$f(\rho) = \frac{1}{2\pi i} \oint_{\Gamma} f(\xi)(\xi - \rho)^{-1} d\xi \quad (3.47)$$

teljesül, ahol a  $\Gamma$  sima görbe pontosan egyszer kerüli meg a  $\rho$  pontot pozitív irányítással és  $\text{Ran } \gamma \subseteq \mathcal{D}$ . Ez a gondolatmenet általánosítása vezet el a Riesz–Dunford-féle operátorkalkulushoz. A témakör bővebben megtalálható a legtöbb terjedelmesebb funkcionálanalízis-könyvben, például Conway [21] könyvében. Ezek szerint az  $A$  önadjungált operátor  $f$  komplex analitikus függvény általi képét a

$$f(A) = \frac{1}{2\pi i} \oint_{\Gamma} f(\xi)(\xi \text{ id} - A)^{-1} d\xi \quad (3.48)$$

kifejezés értelmezi, ahol a  $\Gamma$  sima görbe pontosan egyszer kerüli meg az  $A$  operátor sajátértékeit, pozitív irányítással.



Önadjungált mátrixok esetén ennek az operátorkalkulusnak a bevezetése feleslegesnek tűnhet, azonban a monoton metrikák vizsgálatánál komoly előnyök fognak származni ebből a módszerből.

Láttuk, hogy tetszőleges  $D \in \mathcal{M}_n^+$  állapot esetén az  $L_{n,D}$  és az  $R_{n,D}$  operátor önadjungált, ezért bármilyen komplex analitikus  $f$  függvény esetén alkalmazhatjuk a

$$f(L_{n,D}) = \frac{1}{2\pi i} \oint_{\Gamma} f(\xi)(\xi \text{id} - L_{n,D})^{-1} d\xi \quad f(R_{n,D}) = \frac{1}{2\pi i} \oint_{\Gamma} f(\xi)(\xi \text{id} - R_{n,D})^{-1} d\xi \quad (3.49)$$

képleteket. Az

$$\begin{aligned} (\xi \text{id} - L_{n,D})^{-1}Y = Z &\Rightarrow Y = (\xi \text{id} - L_{n,D})Z = \xi Z - DZ = (\xi \text{id} - D)Z \Rightarrow \\ &\Rightarrow (\xi \text{id} - D)^{-1}Y = Z \end{aligned} \quad (3.50)$$

$$\begin{aligned} (\xi \text{id} - R_{n,D})^{-1}Y = Z &\Rightarrow Y = (\xi \text{id} - R_{n,D})Z = \xi Z - ZD = Z(\xi \text{id} - D) \Rightarrow \\ &\Rightarrow Y(\xi \text{id} - D)^{-1} = Z \end{aligned}$$

egyenletek alapján kapjuk az

$$f(L_{n,D})(X) = \frac{1}{2\pi i} \oint_{\Gamma} f(\xi)(\xi \text{id} - D)^{-1}X d\xi \quad (3.51)$$

$$f(R_{n,D})(X) = \frac{1}{2\pi i} \oint_{\Gamma} f(\xi)X(\xi \text{id} - D)^{-1} d\xi$$

kifejezéseket.

Ezen kifejezések kiterjeszthetők kétváltozós  $(x, y) \mapsto c(x, y)$  komplex analitikus függvényekre is, mivel az  $L_{n,D}$  és az  $R_{n,D}$  operátor felcserélhető [25]. Ekkor a

$$c(L_{n,D}, R_{n,D}) = \frac{1}{(2\pi i)^2} \oint \oint c(\xi, \eta)(\xi \text{id} - L_{n,D})^{-1} \circ (\eta \text{id} - R_{n,D})^{-1} d\xi d\eta \quad (3.52)$$

kifejezést kapjuk, melynek értéke az  $X$  helyen

$$c(L_{n,D}, R_{n,D})(X) = \frac{1}{(2\pi i)^2} \oint \oint c(\xi, \eta)(\xi \text{id} - D)^{-1}X(\eta \text{id} - D)^{-1} d\xi d\eta. \quad (3.53)$$

Ennek a segítségével írhatjuk fel a monoton metrikákat speciális integrál alakjában. Ez nagyon előnyös a differenciálgeometriai vizsgáldások során, mikor szükség lesz a metrikák magasabbrendű deriváltjaira.

Petz tételét a Cencov–Morozova-féle függvényekkel is kifejezhetjük.

**3.14. Tétel.** *Az  $\mathcal{M}_n^+$  sokaságon értelmezett  $K^{(n)}$  monoton Riemann-metrikához, létezik egy  $c_f$  Cencov–Morozova-féle függvény, melyre minden  $D \in \mathcal{M}_n^+$  állapot, és tetszőleges  $X, Y \in T_D\mathcal{M}_n^+$  érintővektorokra*

$$K_D^{(n)}(X, Y) = K_D^{(n),f}(X, Y) = \text{Tr}\left(Xc_f(L_D, R_D)(Y)\right). \quad (3.54)$$



Az  $\mathcal{F}_{[0,\infty]}^{(S,n)}$ -beli függvényekre az alábbi példákat említjük:

$$f_{\text{SM}}(x) = \frac{1+x}{2} \quad (3.55)$$

$$f_{\text{LA}}(x) = \frac{2x}{1+x} \quad (3.56)$$

$$f_{\text{KM}}(x) = \frac{x-1}{\log x} \quad (3.57)$$

$$f_{\text{K1}}(x) = \frac{2(x-1)^2}{(1+x)\log^2 x} \quad (3.58)$$

$$f_{\text{K2}}(x) = \frac{2(x-1)\sqrt{x}}{(1+x)\log x} \quad (3.59)$$

$$f_{\text{P1}}(x) = \frac{2x^{\alpha+1/2}}{1+x^{2\alpha}} \quad 0 \leq \alpha \leq 1/2 \quad (3.60)$$

$$f_{\text{P2}}(x) = \frac{\beta(1-\beta)(x-1)^2}{(x^\beta-1)(x^{1-\beta}-1)} \quad \beta \in [-1, 2] \setminus \{0, 1\}, \quad (3.61)$$

$$f_{\text{WYD}}(x) = \frac{1-\alpha^2}{4} \frac{(x-1)^2}{(1-x^{\frac{1-\alpha}{2}})(1-x^{\frac{1+\alpha}{2}})} \quad \alpha \in [-3, 3] \setminus \{-1, 1\} \quad (3.62)$$

$$f_{\text{WY}}(x) = \frac{1}{4}(\sqrt{x}+1)^2 \quad (3.63)$$

$$f_{\text{P3}}(x) = \left(\frac{1+x^{\frac{1}{\nu}}}{2}\right)^\nu \quad \nu \in [1, 2] \quad (3.64)$$

Az adott  $f \in \mathcal{F}_{[0,\infty]}^{(S,n)}$  függvényhez tartozó  $K^{(n),f}$  metrikát az alábbi gondolatmenettel határozhatjuk meg. A (3.44) képlet segítségével kiszámoljuk a  $c_f$  Cencov–Morozova-féle függvényt. Kihhasználva, hogy a metrika invariáns a  $D$  állapot unitér transzformációival szemben, feltehetjük, hogy az általunk választott  $D \in \mathcal{M}_n^+$  állapot diagonális, főátlójában a  $(\lambda_1, \dots, \lambda_n)$  sajátértékek szerepelnek. A  $K_D^{(n),f}(\cdot, \cdot)$  metrika linearitását kihhasználva, elég meghatározni a metrikának az  $(F_{ij})_{1 \leq i \leq j \leq n}$  és  $(H_{ij})_{1 \leq i < j \leq n}$  mátrixegységeken felvett értékeit. Ezen mennyiség a (3.53) és (3.54) képlet felhasználásával

$$K_D^{(n),f}(X, Y) = \text{Tr} \frac{1}{(2\pi i)^2} \oint \oint c(\xi, \eta) X(\xi \text{id} - D)^{-1} Y(\eta \text{id} - D)^{-1} d\xi d\eta. \quad (3.65)$$

Érdemes most felidézni a 2.6. részben szereplő képleteket. Pontosabban, a metrika kiszámításához kövessük végig a (2.123–2.131) képleteket. A (2.124) kifejezéshez hasonlóan vezessük be a

$$G_f : \mathcal{M}_n^+ \rightarrow \text{Lin}(T_D \mathcal{M}_n^+, T_D \mathcal{M}_n^+) \quad D \mapsto (X \mapsto G_f(D)(X)) \quad (3.66)$$

$$G_f(D)(X) = \frac{1}{(2\pi i)^2} \oint \oint c(\xi, \eta) (\xi - D)^{-1} X (\eta - D)^{-1} d\xi d\eta \quad (3.67)$$

függvényt. (Gyakorlatilag az  $\int_T = \frac{1}{(2\pi i)^2} \oint \oint$  és  $d\mu(t) = c(\xi, \eta) d\xi d\eta$  helyettesítést végeztük el a (2.124) képletten.) A (2.126) kifejezés ekkor így alakul.

$$(G_f(D)(F_{ij}))_{ab} = \frac{1}{(2\pi i)^2} \oint \oint c(\xi, \eta) (\xi - \lambda_a)^{-1} (F_{ij})_{ab} (\eta - \lambda_b)^{-1} d\xi d\eta \quad (3.68)$$

A (2.127) pontban definiált  $m$  függvény ekkor

$$m_f : \mathbb{R}^+ \times \mathbb{R}^+ \rightarrow \mathbb{R} \quad (x, y) \mapsto \frac{1}{(2\pi i)^2} \oint \oint c_f(\xi, \eta) (\xi - x)^{-1} (\eta - y)^{-1} d\xi d\eta. \quad (3.69)$$

Vagyis az

$$m_f(x, y) = c_f(x, y) \quad (3.70)$$

azonosságot kapjuk. Minden  $f \in \mathcal{F}_{[0, \infty]}^{(S, n)}$  függvényre

$$m_f(x, x) = \frac{1}{x} \quad (3.71)$$

adódik.

A 2.24. tétel ekkor így alakul.

**3.15. Tétel.** Tekintsük az  $(\mathcal{M}_n^+, K^{(n), f})$  sokaságot, ahol  $K^{(n), f}$  az  $f \in \mathcal{F}_{[0, \infty]}^{(S, n)}$  függvény által generált monoton metrika. Legyen a  $D \in \mathcal{M}_n^+$  állapot  $D = \sum_{k=1}^n \lambda_k E_{kk}$  alakú.

$$\begin{aligned} \text{Ha } 1 \leq i < j \leq n, \quad 1 \leq k < l \leq n, \text{ akkor:} & \quad \begin{cases} G_D(H_{ij}, H_{kl}) = \delta_{ik} \delta_{jl} 2m(\lambda_i, \lambda_j) \\ G_D(F_{ij}, F_{kl}) = \delta_{ik} \delta_{jl} 2m(\lambda_i, \lambda_j) \\ G_D(H_{ij}, F_{kl}) = 0, \end{cases} \\ \text{ha } 1 \leq i < j \leq n, \quad 1 \leq k \leq n, \text{ akkor:} & \quad G_D(H_{ij}, F_{kk}) = G(F_{ij}, F_{kk}) = 0, \\ \text{ha } 1 \leq i \leq n, \quad 1 \leq k \leq n, \text{ akkor:} & \quad G_D(F_{ii}, F_{kk}) = \delta_{ik} 4m(\lambda_i, \lambda_i) \end{aligned} \quad (3.72)$$

teljesül.

Figyelem: az  $\mathcal{M}_n^+$  tér érintőterében csak nulla nyomú mátrixok vannak, a fenti tétel csak formálisan engedi meg az  $F_{ii}$  alakú diagonális elemek megjelenését.

Érdeemes röviden áttekinteni az  $\mathcal{F}_{[0,\infty]}^{(S,n)}$ -beli függvényekre hozott (3.55–3.64) példák által generált monoton metrikák főbb tulajdonságait.

**3.2. Példa.** Az  $f_{SM}$  (3.55) függvény: A 3.12. tételben szereplő  $\mu$  mértéket  $\frac{\delta_0 + \delta_1}{2}$ -nek választva kapjuk, hogy az  $f_{SM}$  függvény operátor monoton. Sőt igazolható, hogy az  $\mathcal{F}_{[0,\infty]}^{(S,n)}$ -beli függvények halmazában az

$$f \stackrel{\leq}{[0,1]} g \iff f(x) \leq g(x) \quad \forall x \in [0, 1] \quad (3.73)$$

relációra nézve maximális [58]. A metrikát leíró (3.39) kifejezésben azonban  $f^{-1}$  szerepel, ezért az  $f_{SM}$  függvényhez tartozó  $K_{SM}^{(n)}$  metrikát *legkisebb monoton metrikának* (vagy *Bures-metrikának*) nevezzük. Az  $f_{SM}$  függvényhez tartozó Cencov–Morozova-féle függvény

$$c_{SM}(x, y) = \frac{2}{x + y} . \quad (3.74)$$

A (3.14.) tétel alapján, a diagonális  $D \in \mathcal{M}_n^+$  állapot esetén a metrika  $K_{SM,D}^{(n)}(X, Y)$  értékét az alábbi módon határozhatjuk meg. Mivel

$$\begin{aligned} c_{SM}(L_{n,D}, R_{n,D})(Y) = Z &\Rightarrow 2Y = (L_{n,D} + R_{n,D})(Z) \Rightarrow \\ Y &= \frac{DZ + ZD}{2} , \end{aligned} \quad (3.75)$$

ezért a (3.75) egyenletet megoldását adó  $Z$  mátrixot kiszámítva, kapjuk a

$$K_{SM,D}^{(n)}(X, Y) = \text{Tr } XZ \quad (3.76)$$

egyenlőséget. Ezt *szimmetrikus logaritmikus deriválnak* is nevezik.

¶

**3.3. Példa.** Az  $f_{LA}$  (3.56) függvény: A (3.11.) tételben szereplő  $\mu$  mértéket  $\delta_{1/2}$ -nek választva kapjuk, hogy az  $f_{LA}$  függvény operátormonoton és az  $\mathcal{F}_{[0,\infty]}^{(S,n)}$ -beli függvények halmazában a  $\stackrel{\leq}{[0,1]}$  relációra nézve minimális elem [14]. Az  $f_{LA}$  függvényhez tartozó  $K_{LA}^{(n)}$  metrikát *legnagyobb monoton metrikának* nevezzük. Az  $f_{LA}$ -hoz tartozó Cencov–Morozova-féle függvény

$$c_{LA}(x, y) = \frac{x + y}{2xy} . \quad (3.77)$$

A (3.75) egyenletekhez hasonlóan igazolható, hogy adott  $Y \in \mathcal{M}_n$  esetén

$$c_{\text{LA}}(L_{n,D}, R_{n,D})(Y) = \frac{1}{2}(YD^{-1} + D^{-1}Y), \quad (3.78)$$

ezért a nyomképzés ciklikusságát és  $D$  diagonalitását kihasználva kapjuk a

$$K_{\text{LA},D}^{(n)}(X, Y) = \text{Tr } XD^{-1}Y \quad (3.79)$$

egyenlőséget. □

Az  $f_{\text{SM}}$  és  $f_{\text{LA}}$  függvényekre tett megjegyzések alapján kapjuk az alábbi tételt.

**3.16. Tétel.** Minden  $f \in \mathcal{F}_{[0,\infty]}^{(S;n)}$  függvényre

$$f_{\text{SM}_{[0,1]}} \stackrel{\geq}{=} f \stackrel{\geq}{=} f_{\text{LA}} \quad (3.80)$$

teljesül.

**3.4. Példa.** Az  $f_{\text{KM}}$  (3.57) függvény: A 3.12. tételben szereplő  $\mu$  mértéket a Lebesgue-mértéknek választva kapjuk az  $f_{\text{KM}}(x)$  függvény operátormonotonitását. Az  $f_{\text{KM}}$  által generált  $K_{\text{KM}}^{(n)}$  monoton metrikát *Kubo–Mori-metrikának* (vagy *Bogoljubov-féle belső szorzásnak*, vagy *kanonikus korrelációnak*) nevezzük. Az  $f_{\text{KM}}$ -hez tartozó Cencov–Morozova-féle függvény

$$c_{\text{KM}}(x, y) = \frac{\log x - \log y}{x - y}. \quad (3.81)$$

A

$$c_{\text{KM}}(x, y) = \int_0^\infty (t+x)^{-1}(t+y)^{-1} dt \quad (3.82)$$

integrál reprezentációt használva kapjuk, hogy diagonális  $D \in \mathcal{M}_n^+$  esetén

$$K_{\text{KM},D}^{(n)}(X, Y) = \text{Tr} \int_0^\infty X(t \text{id} + L_{n,D})^{-1} Y(t \text{id} + R_{n,D})^{-1} dt = \quad (3.83)$$

$$= \text{Tr} \int_0^\infty X(t + D)^{-1} Y(t + D)^{-1} dt \quad (3.84)$$

teljesül.

A Kubo–Mori-féle belső szorzás központi szerepet játszik a lineárisválasz-elméletben [29, 95]. □

**3.5. Példa.** Az  $f_{P1}$  (3.60) függvény: Petz [82] cikkében igazolta, hogy az  $f_{P1}$  függvény operátormonoton. Úgy érvelt, hogy mivel az

$$f_{P1}^{-1}(x) = \frac{1}{2}x^{-\alpha-1/2} + \frac{1}{2}x^{\alpha-1/2} \quad (3.85)$$

függvény operátor monoton csökkenő, ezért  $f_{P1}$  operátor monoton növény. (Az ehhez szükséges tételek megtalálhatóak például Bhatia [14] könyvében.)

A P1 függvénycsalád folytonos utat jelent az

$$f_{P1}^{(\alpha=0)}(x) = \sqrt{x} \quad f_{P1}^{(\alpha=1/2)}(x) = f_{LA}(x) = \frac{2x}{1+x} \quad (3.86)$$

függvények között. A család tagjaihoz tartozó Cencov–Morozova-féle függvény

$$c_{P1}(x, y) = \frac{x^{2\alpha}y^{2\alpha}}{2x^{\alpha+1/2}y^{\alpha+1/2}}. \quad (3.87)$$

□

**3.6. Példa.** Az  $f_{K1}, f_{K2}$  (3.58, 3.59) függvények: Petz [82] bizonyította ezekről a függvényekről, hogy operátormonotonak. A bizonyítása Kubo és Ando operátorátlagokról szóló [58] cikkében szereplő 6.2. tételen alapul, mely szerint (többek között) a

$$g_{n+1}(x) = \frac{g_n(x) + h_n(x)}{2} \quad h_{n+1}(x) = \sqrt{g_{n+1}(x)h_n(x)} \quad (3.88)$$

kettős sorozat közös határértéke, operátormonoton  $g_1(x)$  és  $h_1(x)$  kezdeti függvények esetén szintén operátormonoton lesz.

A  $g_1(x) = \sqrt{x}$  és  $h_1(x) = \frac{2x}{x+1}$  kezdeti értékekre a közös határérték

$$\frac{x-1}{\log x} \frac{2\sqrt{x}}{1+x}, \quad (3.89)$$

a  $g_1(x) = \frac{x-1}{\log x}$ ,  $h_1(x) = \frac{x-1}{\log x} \frac{2\sqrt{x}}{1+x}$  kezdeti értékekre pedig

$$\left(\frac{x-1}{\log x}\right)^2 \frac{2}{1+x}. \quad (3.90)$$

□

**3.7. Példa.** Az  $f_{P_2}, f_{WYD}$  (3.61, 3.62) függvények: A  $\beta = \frac{1-\alpha}{2}$  helyettesítéssel az  $f_{WYD}$  függvényt vissza lehet vezetni az  $f_{P_2}$  függvényre. Wigner, Yanase és Dyson foglalkozott részletesen az  $f_{WYD}$  függvényből származó Fisher-féle információval, melynek eredete Wigner és Yanase [112] 1963-as cikke. Hasegawa az  $f_{P_2}$  függvény által generált metrikát elemezte [45].

Ennek a monotonitását Hasegawa és Petz igazolta két részletben [46, 86]. A bizonyítás a

$$z^{-\beta} = \frac{\sin \beta \pi}{\pi} \int_0^\infty \frac{\lambda^{-\beta}}{\lambda + z} d\lambda \quad 0 < \beta < 1 \quad (3.91)$$

$$z^{\beta-1} = \frac{\sin \beta \pi}{\pi} \int_0^\infty \frac{\lambda^\beta}{(\lambda + z)^2} d\lambda \quad -1 < \beta < 0$$

egyenlőségeken alapul. Az

$$\frac{1}{f_{P_2}(x)} = \quad (3.92)$$

$$\begin{cases} \frac{\sin \beta \pi}{\pi} \int_0^1 \int_0^1 \int_0^\infty \frac{\lambda^{\beta-1}}{\lambda(x(1-t)+t) + (x(1-s)+s)} d\lambda ds dt & \text{ha } 0 < \beta < 1, \\ \frac{\sin \beta \pi}{\pi} \int_0^1 \int_0^1 \int_0^\infty \frac{\lambda^\beta (x(1-t)+t)}{(\lambda(x(1-t)+t) + (x(1-s)+s))^2} d\lambda ds dt & \text{ha } -1 < \beta < 0, \end{cases}$$

kifejezések segítségével az  $f_{P_2}$  függvény reciprokáról ki lehet mutatni, hogy operátormonoton csökkenő, vagyis az  $f_{P_2}$  függvény operátormonoton növekvő.

Az  $f_{P_2}$  paraméteres függvényt természetes módon ki lehet terjeszteni a  $\beta = 0, 1$  paraméterértékekre.

$$\lim_{\beta \rightarrow 0} f_{P_2}(x) = \lim_{\beta \rightarrow 1} f_{P_2}(x) = f_{KM}(x) = \frac{x-1}{\log x} \quad (3.93)$$

Az  $f_{P_2}$  függvény által generált Cencov–Morozova-féle függvény

$$c_{P_2}(x, y) = \frac{1}{\beta(1-\beta)} \frac{(x^\beta - y^\beta)(x^{1-\beta} - y^{1-\beta})}{y(x-y)^2}. \quad (3.94)$$

¶

**3.8. Példa.** Az  $f_{P_3}$  (3.64) függvény: Az Ando [6] cikkében szereplő 4.3. következmény szerint, ha  $f \in \mathcal{F}_{[0, \infty]}$ , akkor  $\nu \geq 1$  esetén az  $f_\nu(x) = f(x^{1/\nu})^\nu$  függvény is operátormonoton. Elég ezt alkalmazni az  $f_{SM}(x)$  függvényre. ¶

A legkisebb, illetve legnagyobb  $\mathcal{F}_{[0,\infty]}^{(S,n)}$ -beli függvények között a

$$f_{\text{SM}} = f_{\text{P3}}^{(\nu=1)} \underset{[0,1]}{\geq} f_{\text{P3}}^{(1 \leq \nu \leq 2)} \underset{[0,1]}{\geq} f_{\text{P3}}^{(\nu=2)} = f_{\text{WY}} = f_{\text{WYD}}^{(\alpha=0)} \underset{[0,1]}{\geq} f_{\text{WYD}}^{(0 \leq \alpha \leq 3)} \underset{[0,1]}{\geq} f_{\text{WYD}}^{(\alpha=3)} = f_{\text{LA}} \quad (3.95)$$

sorozat folytonos utat határoz meg [37].

A (3.80) egyenlőtlenség segítségével igazolható az alábbi tétel [64].

**3.17. Tétel.** Legyen  $f \in \mathcal{F}_{[0,\infty]}^{(S,n)}$ . Ekkor minden  $D \in \mathcal{M}_n^+$  állapotra és  $X \in T_D \mathcal{M}_n^+$  érintővektorra fennállnak az

$$K_{LA,D}^{(n)}(X, X) \geq K_D^{(n),f}(X, X) \geq K_{SM,D}^{(n)}(X, X) \quad (3.96)$$

egyenlőtlenségek.

A klasszikus esetben a Cramer–Rao-egyenlőtlenség mutatta meg, hogy a Fisher-féle információ bizonyos határt jelent a torzítatlan becslések szórására. Kvantumos esetben egyelőre csak az egyparaméteres statisztikai modellre definiáltuk a kvantumos Fisher-féle információt, azonban az messze nem volt egyértelmű. Az alábbiakban a Cramer–Rao-egyenlőtlenség nemkommutatív kiterjesztését mutatjuk be.

**3.15. Definíció.** Legyen  $(L_n, \mathcal{Q}, \Xi)$   $m$ -dimenziós kvantummechanikai statisztikai modell és  $f \in \mathcal{F}_{[0,\infty]}^{(S,n)}$  tetszőleges függvény. A  $D_{\vartheta_0}$  pontban a *kvantummechanikai Fisher-féle információs mátrixot* az

$$\left( I_{D_{\vartheta_0}}^f \right)_{ij} = \text{Tr} \left[ \left( \frac{\partial D_{\vartheta}}{\partial \vartheta_i} \Big|_{\vartheta=\vartheta_0} \right) \left( R_{n,D_{\vartheta_0}}^{1/2} f(L_{n,D_{\vartheta_0}} R_{n,D_{\vartheta_0}}^{-1}) R_{n,D_{\vartheta_0}}^{1/2} \right)^{-1} \left( \frac{\partial D_{\vartheta}}{\partial \vartheta_j} \Big|_{\vartheta=\vartheta_0} \right) \right] \quad (3.97)$$

kifejezés definiálja  $(i, j = 1, \dots, m)$ .

A Fisher-féle információs mátrix  $(i, j)$ -edik eleme kifejezhető az  $f$  által generált Riemann-metrikával

$$\left( I_{D_{\vartheta_0}}^f \right)_{ij} = K_{D_{\vartheta_0}}^{(n),f} \left( \frac{\partial D_{\vartheta}}{\partial \vartheta_i} \Big|_{\vartheta=\vartheta_0}, \frac{\partial D_{\vartheta}}{\partial \vartheta_j} \Big|_{\vartheta=\vartheta_0} \right). \quad (3.98)$$

A kvantummechanikai Fisher-féle információra szintén teljesül egyfajta monotonitási tulajdonság [3].

**3.18. Tétel.** Legyen  $(L_n, \mathcal{Q}_1, \Xi)$  kvantummechanikai statisztikai modell és legyen  $\alpha : \mathcal{Q}_1 \rightarrow \mathcal{M}_n^+$  sztochasztikus leképezés. Definiáljuk az  $(L_n, \mathcal{Q}_2, \Xi)$  statisztikai modellt, mint a  $\mathcal{Q}_1$  halmaz  $\alpha$  általi képét. (Vagyis használjuk az

$$i_2 : \Xi \rightarrow \mathcal{Q}_2 \quad \vartheta \mapsto \alpha(D_\vartheta) \quad (3.99)$$

paraméterezést.) Tetszőleges  $f \in \mathcal{F}_{[0,\infty]}^{(S,n)}$  függvény esetén jelölje  $I_{D,1}^f$  és  $I_{\alpha(D),2}^f$  a  $\mathcal{Q}_1$  illetve  $\mathcal{Q}_2$  modell  $D$  illetve  $\alpha(D)$  pontbeli kvantummechanikai Fisher-féle információs mátrixát. Ekkor az

$$I_{\alpha(D),2}^f \leq I_{D,1}^f \quad (3.100)$$

egyenlőtlenség teljesül minden  $D \in \mathcal{Q}_1$  állapotra.

A Cramer-Rao-egyenlőtlenség megfogalmazásához szükséges a nemkommutatív paraméterbecslés értelmezése. Egyelőre azonban csak olyan becslésekkel fogunk számolni, melyek bizonyos fizikai mennyiségek mérésének felelnek meg.

**3.16. Definíció.** Legyen  $(L_n, \mathcal{Q}, \Xi)$   $m$ -dimenziós kvantummechanikai modell. A valós (illetve komplex) állapot tér esetén, a valós (illetve komplex) elemeket tartalmazó  $n \times n$ -es, önadjungált operátorok tetszőleges  $A_1, \dots, A_m$  halmazát *mérésből adódó paraméterbecslésnek* nevezzük.

Azt mondjuk, hogy a becslés *torzítatlan*, ha minden  $\vartheta \in \Xi$  esetén

$$\vartheta = (\vartheta_1, \dots, \vartheta_m) = (\text{Tr } D_\vartheta A_1, \dots, \text{Tr } D_\vartheta A_m) \quad (3.101)$$

teljesül.

Adott  $f \in \mathcal{F}_{[0,\infty]}^{(S,n)}$  függvényre a becslés  $f$ -varianciája a  $D_\vartheta$  pontban a

$$\left( V_{D_\vartheta}^{(n),f}(A_1, \dots, A_m) \right)_{ij} = K_{D_\vartheta}^{(n),f}(A_i - \vartheta_i, A_j - \vartheta_j) \quad i, j = 1, \dots, m \quad (3.102)$$

mátrix.

Az állapot tér geometriai struktúrájának a vizsgálata során először Petz [81], valamint Petz és Tóth [89] fogalmazta meg az alábbi nemkommutatív Cramer-Rao egyenlőtlenség egy speciális, a Kubo-Mori-skalárszorzatra vonatkozó, esetét. A mostani formája tetszőleges monoton metrikát megenged [84].

**3.19. Tétel.** (Nemkommutatív Cramer-Rao-egyenlőtlenség:) Legyen  $(L_n, \mathcal{Q}, \Xi)$   $m$ -dimenziós kvantummechanikai statisztikai modell,  $A_1, \dots, A_m$  mérésből adódó torzítatlan paraméterbecslés és  $f \in \mathcal{F}_{[0,\infty]}^{(S,n)}$  tetszőleges függvény. Ekkor minden  $D_\vartheta \in \mathcal{Q}$  állapotra

$$V_{D_\vartheta}^{(n),f}(A_1, \dots, A_m) \geq \left( I_{D_\vartheta}^f \right)^{-1} \quad (3.103)$$

teljesül.



A klasszikus esetben megszokott Cramer–Rao-tételből, most az  $f$  operátormonoton függvény miatt „sok” tétel lett. Nehéz megmondani általánosan, hogy mely  $f$  függvénynél tartalmazza a legtöbb információt a tétel; ugyanis adott statisztikai sokaság esetén az  $f$ -től függő Fisher-információk inverzét könnyen össze lehet hasonlítani, azonban ezzel párhuzamosan változik a variancia is.

A fenti tétel gyakorlati alkalmazhatóságával komoly problémák vannak. Ezek közül néhányat megemlítünk, de részletesen egyikkel sem foglalkozunk.

Az egyik fő problémát a mérés jelenti. A mérés folyamán a mérőeszköz megváltoztatja a mérendő fizikai objektumot. Ezen változás leírása sok absztrakt fizikai modellnek része (pl. a Neumann-féle kvantummechanikai axiómákon alapuló modell), azonban a kérdés még nincs kielégítően megoldva. Ahhoz, hogy a variancia kézzelfogható mennyiség legyen nem egy, hanem több mérést kell elvégezni. A második mérés azonban már nem az eredeti rendszeren történik, hiszen az első mérés már megváltoztatta azt. Továbbá még nincs megoldva az a probléma sem, hogy a különböző operátormonoton függvények, illetve a belőlük származtatott mennyiségek (Riemann-metrika, Fisher-információ, a később bevezetendő relatív entrópia), milyen fizikai tartalommal bírnak.

A nemkommutatív paraméterbecslés különböző megközelítési módjai, valamint általános tulajdonságai megtalálhatók Nagaoka [35, 36], Matsumoto és Hayashi [47, 48] munkáiban.

Térjünk vissza az előző részben (a 3.3. pontban) bevezetett információs mennyiségek és a Fisher-metrika kapcsolatára. Legyen  $f \in \mathcal{F}_{[0,\infty]}^{(S)}$ ,  $D \in \mathcal{M}_n^+$  és legyen  $X, Y \in T_D \mathcal{M}_n^+$ . Ha  $X$  vagy  $Y$  felcserélhető  $D$ -vel, akkor

$$K_D^{(n),f}(X, Y) = \frac{1}{f(1)} \operatorname{Tr} X D^{-1} Y \quad (3.104)$$

teljesül. Tehát ha vizsgálódásainkat megszorítjuk a kommutatív (pl. diagonális) esetre, akkor szintén azt kapjuk, hogy a Fisher-féle információs mátrix számszorzó erejéig egyértelmű.

Adott  $D \in \mathcal{M}_n^+$  pontban a  $T_D \mathcal{M}_n^+$  érintőtérben értelmezzük az

$$T_D^C \mathcal{M}_n^+ = \{X \in T_D \mathcal{M}_n^+ \mid [X, D] = 0\} \quad (3.105)$$

alteret, illetve a Hilbert-Schmidt skalárszorzásra nézve rá merőleges  $T_D^\perp \mathcal{M}_n^+$  alteret. Igazolható, hogy ekkor minden  $X \in T_D \mathcal{M}_n^+$  érintővektor felírható

$$X = X^C + i[D, X^K] \quad (3.106)$$

alakban, ahol  $X^C \in T_D^C \mathcal{M}_n^+$ ,  $i[D, X^K] \in T_D^\perp \mathcal{M}_n^+$  és  $X^K$  önadjungált.

A (1.5.) példában láttuk, hogy a klasszikus esetben az entrópia második deriváltjából is lehetett származtatni a Fisher-féle információt. Az alábbi példából kiderül, hogy a Kubo–Mori-féle metrika származtatható a Neumann-féle entrópia második deriváltjából. A Kubo–Mori-féle metrika ezen tulajdonságának a fizikai alkalmazása Balian, Alhassid és Reinhardt [9] cikkében található.

**3.9. Példa.** A log függvényre az

$$\log x = \int_0^\infty (1+t)^{-1} - (x+t)^{-1} dt \quad (3.107)$$

integrálrepresentációt használva az entrópia az

$$S(D) = \text{Tr } D \int_0^\infty (D+t)^{-1} - (\text{id}+t)^{-1} dt \quad (3.108)$$

alakban is felírható. Az  $S(D+A) - S(D)$  kifejezésben az  $A$ -tól elsőrendben függő rész meghatározásához vezessük be az  $f(A) \cong g(A)$  jelölést, melynek jelentése, hogy  $\|f(A) - g(A)\| = o(\|A\|^2)$ , azaz  $f$  és  $g$  ugyanúgy függ  $A$ -tól nullad és első rendben, vagyis

$$\lim_{A \rightarrow 0} \frac{f(A) - g(A)}{\|A\|} = 0 \quad (3.109)$$

teljesül.

$$\begin{aligned} S(D+A) - S(D) &= \text{Tr}(D+A) \int_0^\infty (D+A+t)^{-1} - (1+t)^{-1} dt - S(D) \cong \\ &\cong \text{Tr}(D+A) \int_0^\infty (D+t)^{-1} - (D+t)^{-1}A(D+t)^{-1} - (1+t)^{-1} dt - S(D) = \\ &= \text{Tr } D \int_0^\infty (D+t)^{-1} - (D+t)^{-1}A(D+t)^{-1} - (1+t)^{-1} dt + \\ &\quad + A \int_0^\infty (D+t)^{-1} - (1+t)^{-1} dt - S(D) = \\ &= \text{Tr } D \int_0^\infty -(D+t)^{-1}A(D+t)^{-1} dt + A \int_0^\infty (D+t)^{-1} - (1+t)^{-1} dt = \\ &= -\text{Tr } A \int_0^\infty (D+t)^{-1}D(D+t)^{-1} dt + A \int_0^\infty (D+t)^{-1} - (1+t)^{-1} dt = \\ &= -\sum_{i,j=1}^n A_{ij} \int_0^\infty \frac{\lambda_i}{(\lambda_i+t)^2} \delta_{ij} dt + A \int_0^\infty (D+t)^{-1} - (1+t)^{-1} dt = \\ &= -\sum_{i=1}^n A_{ii} + A \int_0^\infty (D+t)^{-1} - (1+t)^{-1} dt = -A \log D \end{aligned} \quad (3.110)$$

Mivel az érintőtérbeli  $A \in T\mathcal{M}_n$  vektorokra  $\text{Tr } A = 0$  teljesül, ezért az entrópia deriváltját a fenti számolások alapján a

$$dS : \mathcal{M}_n^+ \rightarrow \text{Lin}(T\mathcal{M}_n, \mathbb{R}) \quad D \mapsto \left( A \mapsto dS(D)(A) \right) \quad (3.111)$$

$$dS(D)(A) = -\text{Tr } A \log D$$

kifejezés adja meg. Az előzőhöz hasonlóan határozhatjuk meg az entrópia második deriváltját: meghatározzuk a  $dS(D+B)(A) - dS(D)(A)$  kifejezésben a  $B$ -től elsőrendben függő tagokat.

$$dS(D+B)(A) - dS(D)(A) = \text{Tr } A \int_0^\infty (D+B+t)^{-1} - (1+t)^{-1} dt - dS(D)(A) \cong \quad (3.112)$$

$$\cong \text{Tr } A \int_0^\infty (D+t)^{-1} - (D+t)^{-1}B(D+t)^{-1} - (1+t)^{-1} dt - dS(D)(A) =$$

$$= \text{Tr } A \int_0^\infty -(D+t)^{-1}B(D+t)^{-1} dt$$

Vagyis az entrópia második deriváltja

$$d^2S : \mathcal{M}_n^+ \rightarrow \text{Lin}\left(T\mathcal{M}_n, \text{Lin}(T\mathcal{M}_n, \mathbb{R})\right) \quad D \mapsto \left( A \mapsto (B \mapsto d^2S(D)(A)(B)) \right) \quad (3.113)$$

$$d^2S(D)(A)(B) = -\text{Tr} \int_0^\infty (D+t)^{-1}A(D+t)^{-1}B dt .$$

Ennek a  $(-1)$ -szereséből is származtatható a Kubo–Mori-féle metrika.

$$\langle \cdot, \cdot \rangle_D^{(\text{KM})} : \mathcal{M}_n^+ \rightarrow \text{Lin}(T\mathcal{M}_n \times T\mathcal{M}_n, \mathbb{R}) \quad D \mapsto \langle \cdot, \cdot \rangle_D^{(\text{KM})} \quad (3.114)$$

$$\langle A, B \rangle_D^{(\text{KM})} = \text{Tr} \int_0^\infty (D+t)^{-1}A(D+t)^{-1}B dt$$

¶

A klasszikus esetben a Fisher-metrikát úgy is meg lehet kapni, mint egy kellően sok dimenziós euklideszi térben lévő gömb euklideszi metrikájának a visszahúzottját. (Lásd a 2.3. példát.) Kvantumos esetben Giblisco és Isola megmutatta, hogy a megfelelő metrika az  $f_{\text{WY}}$  operátormonoton függvényhez (pontosabban pozitív számszorosáshoz) tartozó metrika lesz [38].

**3.10. Példa.** Komplex illetve valós  $\mathcal{M}_n^+$  állapotter esetén jelölje  $S_1^{n^2-1}$  a  $\mathbb{C}^{n \times n}$  illetve az  $\mathbb{R}^{n \times n}$  térben lévő egységsugarú gömböt. Definiáljuk a

$$\phi : \mathcal{M}_n^+ \rightarrow S^{n^2-1} \quad D \mapsto \phi(D) = \sqrt{D} \quad (3.115)$$

leképezést. Minden  $D \in \mathcal{M}_n^+$  esetén a  $\phi(D)\phi(D) = D$  egyenlőségből, a Leibniz-szabály alapján

$$\left( (d_D\phi)(A) \right) \sqrt{D} + \sqrt{D} \left( (d_D\phi)(A) \right) = A \quad \forall A \in T_D\mathcal{M}_n^+ \quad (3.116)$$

adódik. Ez alapján a  $\phi$  leképezés deriváltja a  $D \in \mathcal{M}_n^+$  pontban

$$(d_D\phi)(A) = \left( L_D^{1/2} + R_D^{1/2} \right)^{-1} (A). \quad (3.117)$$

Annak igazolásához, hogy az  $(L_D^{1/2} + R_D^{1/2})^{-1}$  operátor önadjungált, tetszőleges  $X, Y \in T_D\mathcal{M}_n^+$  vektorok esetén vezessük be az

$$a = \langle (L_D^{1/2} + R_D^{1/2})^{-1}(X), Y \rangle \quad b = \langle X, (L_D^{1/2} + R_D^{1/2})^{-1}(Y) \rangle \quad (3.118)$$

$$A = (L_D^{1/2} + R_D^{1/2})^{-1}(X) \quad B = (L_D^{1/2} + R_D^{1/2})^{-1}(Y)$$

jelöléseket. Ekkor  $X = D^{1/2}A + AD^{1/2}$  és  $Y = D^{1/2}B + BD^{1/2}$  teljesül. Az

$$a = \text{Tr}(AY) = \text{Tr}(AD^{1/2}B + ABD^{1/2}) \quad b = \text{Tr}(XB) = \text{Tr}(D^{1/2}AB + AD^{1/2}B) \quad (3.119)$$

egyenletek pedig igazolják, hogy  $a = b$ , vagyis az  $(L_D^{1/2} + R_D^{1/2})^{-1}$  operátor önadjungált.

Az  $S_1^{n^2-1}$  gömbön értelmezett Riemann-metrika  $\phi$ -általi visszahúzottjának az értéke tetszőleges  $A, B \in T_D\mathcal{M}_n^+$  vektorokon az alábbi.

$$(\phi^*g)(A, B) = \langle (d_D\phi)(A), (d_D\phi)(B) \rangle = \langle (L_D^{1/2} + R_D^{1/2})^{-1}(A), (L_D^{1/2} + R_D^{1/2})^{-1}(B) \rangle = \quad (3.120)$$

$$= \langle A, (L_D^{1/2} + R_D^{1/2})^{-2}(B) \rangle = \text{Tr} A (L_D^{1/2} + R_D^{1/2})^{-2}(B) =$$

$$= \frac{1}{4} \text{Tr} A c_{\text{WY}}(L_D, R_D)(B)$$

□

Ezek alapján az  $(\mathcal{M}_n^+, K_{\text{WY}}^{(n)})$  Riemann-térben egyszerűen meg lehet határozni a geodetikusokat.

**3.11. Példa.** A (2.5.) példa alapján igazolható, hogy az  $(\mathcal{M}_n^+, K_{\text{WY}}^{(n)})$  Riemann-térben a  $D_1, D_2$  állapotokat összekötő geodetikus

$$\gamma : [0, t_0] \rightarrow \mathcal{M}_n^+ \quad t \mapsto D_t = \left( \frac{D_2^{1/2} - D_1^{1/2} \text{Tr}(D_1^{1/2} D_2^{1/2})}{\sqrt{1 - \text{Tr}(D_1^{1/2} D_2^{1/2})^2}} \sin \frac{t}{2} + D_1^{1/2} \cos \frac{t}{2} \right)^2, \quad (3.121)$$

ahol

$$t_0 = 2 \arccos \text{Tr}(D_1^{1/2}, D_2^{1/2}). \quad (3.122)$$

Vagyis a  $D_1$  és  $D_2$  pontok geodetikus távolsága

$$d_{\text{WY}}(D_1, D_2) = 2 \arccos \text{Tr} D_1^{1/2} D_2^{1/2}. \quad (3.123)$$

□

Az előző példák alapján felmerül a kérdés, hogy mely metrikák állnak elő az euklideszi metrika visszahúzottjaként. Ezt a kérdést Giblisco és Isola [38] tisztázták 2003-ban, azonban erre most nem térünk ki.

### 3.6. Relatív entrópia

A relatív entrópiát kétféle megközelítéssel lehet kiterjeszteni a nemkommutatív esetre. Az egyik az implicit definíció, mikor felsorolunk bizonyos tulajdonságokat és az azokat teljesítő objektumokat relatív entrópiának nevezzük; a másik az explicit definíció, mikor konkrét konstrukció segítségével definiáljuk a relatív entrópiát.

Most az első megközelítési módot mutatjuk be. A klasszikus esetben bevezetett általánosított divergencia (vagy kontrasztfüggvény) főbb, könnyen általánosítható jellemzőit követeljük meg a nemkommutatív relatív entrópiától.

**3.17. Definíció.** (Ruskai [64]:) Egy

$$H(\cdot, \cdot) : \mathcal{M}_n^+ \times \mathcal{M}_n^+ \rightarrow \mathbb{R} \quad (D_1, D_2) \mapsto H(D_1, D_2) \quad (3.124)$$

függvényt *relatív entropikus távolságnak* nevezzük, ha minden  $D_1, D_2 \in \mathcal{M}_n^+$  állapot esetén

1.  $H(D_1, D_2) \geq 0$ , valamint  $H(D_1, D_2) = 0$  pontosan akkor teljesül, ha  $D_1 = D_2$ ,
2. a  $H$  függvény konvex a két változójában, azaz minden  $D_1, D_2, D_3, D_4$  állapot és  $\lambda \in [0, 1]$  paraméter esetén

$$H(\lambda D_1 + (1 - \lambda) D_2, \lambda D_3 + (1 - \lambda) D_4) \leq \lambda H(D_1, D_3) + (1 - \lambda) H(D_2, D_4) \quad (3.125)$$

teljesül.

Azt mondjuk, hogy a  $H$  relatív entropikus távolság

- *monoton*, ha minden  $T : \mathcal{M}_n^+ \rightarrow \mathcal{M}_n^+$  sztochasztikus leképezés esetén

$$H(T(D_1), T(D_2)) \leq H(D_1, D_2) \quad \forall D_1, D_2 \in \mathcal{M}_n^+, \quad (3.126)$$

- *szimmetrikus*, ha  $H(D_1, D_2) = H(D_2, D_1)$ ,
- *differenciálható*, ha minden  $D_1, D_2 \in \mathcal{M}_n^+$  és  $A \in T_{D_1}\mathcal{M}_n^+$ ,  $B \in T_{D_2}\mathcal{M}_n^+$  érintővektorra a

$$h_{D_1, D_2}^{A, B} : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R} \quad (x, y) \mapsto h_{D_1, D_2}^{A, B}(x, y) = H(D_1 + xA, D_2 + yB) \quad (3.127)$$

függvény differenciálható a  $(0, 0)$  pontban.

A relatív entrópia általánosításának (explicit) megközelítését kapjuk, ha a klaszikus esetben definiált Csiszár-féle  $f$ -divergenciákat kiterjesztjük a nemkommutatív esetre. A Csiszár-féle entrópia definíciójában szereplő konvex függvények szerepét az operátorkonvex függvények veszik át.

**3.18. Definíció.** Azt mondjuk, hogy az  $f : \mathbb{R} \rightarrow \mathbb{R}$  folytonos függvény *operátorkonvex*, ha minden  $n \in \mathbb{N}$  számra, minden  $A, B$  önadjungált  $n \times n$ -es mátrixra és minden  $\lambda \in [0, 1]$  paraméterre

$$f(\lambda A + (1 - \lambda)B) \leq \lambda f(A) + (1 - \lambda)f(B) \quad (3.128)$$

teljesül.

Ezen függvények segítségével lehet általánosítani a Csiszár-féle divergenciákat. A kvantumumos általánosítás nehézsége abban rejlik, hogy a  $g(x/y)$  mennyiségben szereplő hányadost többféle módon lehet értelmezni nemkommutatív  $x$  és  $y$  változók esetén.

A legelső lépések közé tartozik Umegaki 1962-es [107] publikációja, ahol az  $D_1, D_2 \in \mathcal{M}_n^+$  állapotok relatív entrópiáját definiálta az

$$S(D_1, D_2) = \text{Tr } D_1(\log D_1 - \log D_2) \quad (3.129)$$

képlettel. Ennek a matematikai fizikában való használatát Lindblad vizsgálta a 70-es években [68, 66, 67]. A relatív entrópia matematikai tulajdonságait többek között Uhlmann [104] és Petz [79] elemezték. A Csiszár-féle divergenciák nemkommutatív általánosításával kapcsolatos Ruskai 1994-es eredménye [94], ahol az entrópia kontrakciós együtthatót próbálta nemkommutatív esetre általánosítani, azonban számottevően új eredményt nem ért el. Következő lépésben Petz és Ruskai [87] 1998-ban a  $P/Q$  „operátorhányadost” a  $P^{-1/2}QP^{-1/2}$  kifejezéssel értelmezték. Az így bevezetett relatív

entrópiára számos tételt bizonyítottak sikerrel. Ezen kiterjesztéseknél azonban sikere-  
sebbnek bizonyult az Araki által bevezetett relatív moduláris operátor használata. Ezt  
a megközelítési módot a relatív entrópiával kapcsolatban Petz már korábban használta  
[77, 79, 83]. Az általánosításnak ezt a módját követjük a továbbiakban. (A relatív  
entrópia operátoralgebrákra is kiterjeszthető. Ez a kiterjesztés és a relatív entrópia  
tulajdonságai megtalálhatóak Lindblad [68], Uhlmann [104], Ohya és Petz [77, 80]  
munkáiban.)

**3.19. Definíció.** (Petz [79]:) Legyen  $g : ]0, \infty] \rightarrow \mathbb{R}$  olyan operátorkonvex függvény,  
melyre  $g(1) = 0$  teljesül. Ekkor a

$$H_g(\cdot, \cdot) : \mathcal{M}_n^+ \times \mathcal{M}_n^+ \rightarrow \mathbb{R} \quad (D_1, D_2) \mapsto H_g(D_1, D_2) = \text{Tr} \left( D_1^{1/2} g(L_{D_2} R_{D_1}^{-1}) D_1^{1/2} \right) \quad (3.130)$$

függvényt  $g$ -relatív entrópiának nevezzük.

Az operátormonoton függvényekhez hasonlóan, a  $g$  operátorkonvex függvény  
*transzponáltja* legyen  $g^\lambda(x) = xg(x^{-1})$ , *duálisa* pedig  $g^\perp(x) = \frac{x}{g(x)}$ . A  $g$  operátorkonvex  
függvényről azt mondjuk, hogy *szimmetrikus*, ha  $g^\lambda = g$  teljesül. Adott  $I \subseteq \mathbb{R}$   
intervallumra jelölje  $\mathcal{K}_I$  azon  $g : I \rightarrow \mathbb{R}$  operátorkonvex függvények halmazát, melyekre  
 $1 \in I$  esetén  $g(1) = 0$  teljesül. Azt mondjuk, hogy a  $g \in \mathcal{K}_{]0, \infty]}$  függvény *normált*, ha  
 $g''(1) = 1$ . Az  $I$  halmazon értelmezett, normált operátorkonvex függvények halmazát  
jelölje  $\mathcal{K}_I^{(n)}$ . A  $\mathcal{K}_I$ - illetve  $\mathcal{K}_I^{(n)}$ -beli szimmetrikus elemek halmazát jelölje  $\mathcal{K}_I^{(S)}$  illetve  
 $\mathcal{K}_I^{(S, n)}$ .

**3.12. Példa.** Adott  $D_1, D_2 \in \mathcal{M}_n^+$  állapotok és  $g \in \mathcal{K}_{]0, \infty]}$  függvény esetén  $H_g(D_1, D_2)$   
értékét az alábbi módon határozhatjuk meg. Tekintsük az állapotok

$$D_1 = \sum_{i=1}^n \mu_i P_i \quad D_2 = \sum_{i=1}^n \nu_i Q_i \quad (3.131)$$

spektrális felbontását, ahol a  $(P_i)_{i=1, \dots, n}$  illetve  $(Q_i)_{i=1, \dots, n}$  projekciók rendszere páron-  
ként ortogonális és rangjuk egy. Ekkor a  $g$ -relatív entrópiára

$$D_g(D_1, D_2) = \sum_{i, j=1}^n \mu_i g \left( \frac{\nu_j}{\mu_i} \right) \text{Tr} P_i Q_j \quad (3.132)$$

teljesül. ¶

A klasszikus esetet bemutató 1.3. részben láttuk, hogy a Csiszár-féle  $f$ -divergencia,  
alkalmasan megválasztott  $f$ -függvény esetén visszaadja a főbb divergenciatípusokat. A

nemkommutatív esetben a fordított utat járjuk be. A klasszikus esetben jól használható függvényekkel definiáljuk a klasszikus divergenciák nemkommutatív megfelelőit.

A  $g \in \mathcal{K}_{[0,\infty]}$  függvény transzponálásának a hatása a relatív entrópiára az, hogy megcseréli az argumentumokat, vagyis

$$H_g(D_1, D_2) = H_{g\setminus}(D_2, D_1) . \quad (3.133)$$

A  $H_g$   $g$ -relatív entrópia általában nem szimmetrikus, azonban a

$$H_g^{(\text{szim})}(D_1, D_2) = \frac{H_g(D_1, D_2) + H_g(D_2, D_1)}{2} \quad (3.134)$$

képlettel értelmezhetjük a *szimmetrizáltját*, melyet a

$$H_g^{(\text{szim})}(D_1, D_2) = H_{(g+g\setminus)/2}(D_1, D_2) \quad (3.135)$$

formában is írhatunk.

**3.20. Tétel.** [79] Minden  $g$ -relatív entrópia relatív entropikus távolság.

A tétel megfordítása azonban nem igaz. Például a Petz és Ruskai által használt

$$\tilde{H}(D_1, D_2) = \text{Tr } D_1 \log(D_1^{-1/2} D_2 D_1^{-1/2}) \quad (3.136)$$

kifejezés relatív entropikus távolság, azonban nem  $g$ -relatív entrópia [87].

A fenti (3.136) relatív entrópiához formálisan hasonlóan definiált 1982-ben Belavkin. A

$$\tilde{H}(D_1, D_2) = \text{Tr } D_2 \log(D_2^{1/2} D_0^{-1} D_2^{1/2}) \quad (3.137)$$

kifejezést tekintette a relatív entrópia általánosításának [12].

**3.21. Tétel.** [79] Minden  $g$ -relatív entrópia monoton és differenciálható relatív entropikus távolság.

A  $g$ -relatív entrópia monotonitásának a bizonyítása megtalálható Petz [77, 79, 85, 87] munkáiban.

**3.22. Tétel.** Legyen  $g \in \mathcal{K}_{[0,\infty]}$ . Ekkor létezik  $a_g \in \mathbb{R}$ ,  $b_g, c_g \in [0, \infty]$  paraméter és  $\mu_g$  pozitív mérték az  $[0, \infty]$  intervallumon, a  $\mu_g([0, \infty]) < \infty$  feltétellel, úgy, hogy

$$g(x) = a_g(x-1) + b_g(x-1)^2 + c_g \frac{(x-1)^2}{x} + \int_0^\infty (x-1)^2 \frac{1+t}{x+t} d\mu_g(t) \quad (3.138)$$

teljesül. Valamint minden  $a \in \mathbb{R}$ ,  $b, c \in [0, \infty]$  paraméterre, és minden olyan  $[0, \infty]$  halmazon értelmezett pozitív  $\mu$  mértékre, melyre  $\mu([0, \infty]) < \infty$  teljesül, a (3.138) kifejezés operátorkonvex függvényt értelmez.



A  $H_g(D_1, D_2)$  értéke főként a  $D_1 - D_2$  különbségtől függ.

**3.23. Tétel.** [64] Legyen  $g \in \mathcal{K}_{[0, \infty]}$ , ekkor minden  $D_1, D_2 \in \mathcal{M}_n^+$  állapotra

$$H_g(D_1, D_2) = \text{Tr} \left( (D_1 - D_2) R_{D_1}^{-1} \left( g(L_{D_2} R_{D_1}^{-1})(D_1 - D_2) \right) \right) \quad (3.139)$$

teljesül.

Ingarden mutatta meg 1982-ben, hogy bizonyos alaptulajdonságokkal rendelkező relatív entrópiák segítségével lehet Riemann-metrikát indukálni az állapotok terén [53]. A  $g$ -relatív entrópiák mind rendelkeznek ezekkel az alaptulajdonságokkal.

**3.24. Tétel.** Legyen  $g \in \mathcal{K}_{[0, \infty]}$ . Ekkor a

$$K^{g, (n)} : \mathcal{M}_n^+ \rightarrow \text{Lin}(T\mathcal{M}_n \times T\mathcal{M}_n, \mathbb{R}) \quad D \mapsto \left( (X, Y) \mapsto K_D^{g, (n)}(X, Y) \right) \quad (3.140)$$

$$K_D^{g, (n)}(X, Y) = - \frac{\partial^2}{\partial s \partial t} H_g(D + tX, D + sY) \Big|_{t=s=0} \quad (3.141)$$

kifejezés Riemann-metrikát értelmez az  $\mathcal{M}_n^+$  sokaságon. Továbbá a

$$K_D^{g, (n)}(X, Y) = \text{Tr} X G_D^g(Y) \quad \forall D \in \mathcal{M}_n^+, \forall X, Y \in T\mathcal{M}_n \quad (3.142)$$

kifejezés egyértelműen definiál egy

$$G^g : \mathcal{M}_n^+ \rightarrow (T\mathcal{M}_n \rightarrow T\mathcal{M}_n) \quad D \mapsto \left( X \mapsto G_D^g(X) \right) \quad (3.143)$$

függvényt.

Az így generált metrika tulajdonságait először Hasegawa vizsgálta [45]. A metrika monotonitását Hasegawa és Petz bizonyította be [46, 86].

Figyelem, a  $K^{(n), h}$  kifejezést  $h \in \mathcal{F}_{[0, \infty]}$  függvényekre értelmeztük a (3.39) képlettel, míg a  $K^{h, (n)}$  kifejezést  $h \in \mathcal{K}_{[0, \infty]}$  függvényekre a (3.140) képlettel.

Két kérdés merül fel: különböző  $g$  függvények különböző metrikát definiálnak-e, valamint, hogy mi a kapcsolat az így definiált  $K^{g, (n)}$  metrika és a már korábban definiált  $K^{(n), f}$  metrika között.

A  $\mathcal{K}_{[0, \infty]}$  halmaz elemei között vezessük be az  $\sim$  relációt.

$$f \sim g \iff f + f^\setminus = g + g^\setminus \quad (3.144)$$

A  $\sim$  reláció ekvivalenciareláció.

A relatív entrópiák és monoton metrikák közötti összefüggésre vonatkozó alábbi négy tétel bizonyítása megtalálható Lesniewski és Ruskai [64] publikációjában.

**3.25. Tétel.** A  $g_1, g_2 \in \mathcal{K}_{[0,\infty]}$  függvények pontosan akkor generálnak azonos metrikát a (3.140) képlet segítségével, ha  $g_1 \sim g_2$  teljesül.

**3.26. Tétel.** A

$$\phi : \mathcal{K}_{[0,\infty]}^S \rightarrow \mathcal{F}_{[0,\infty]}^S \quad g(x) \mapsto \phi(g)(x) = \begin{cases} \frac{(x-1)^2}{g(x) + xg(x^{-1})} & \text{ha } x > 0, x \neq 1, \\ \frac{1}{g''(1)} & \text{ha } x = 1, \\ \frac{1}{\lim_{x \rightarrow 0} g(x) + xg(x^{-1})} & \text{ha } x = 0 \end{cases} \quad (3.145)$$

leképezés jól értelmezett, továbbá  $K^{g,(n)} = K^{(n),\phi(g)}$  teljesül, azaz

$$K_D^{g,(n)}(X, Y) = K_D^{(n),\phi(g)}(X, Y) \quad \forall D \in \mathcal{M}_n^+ \quad \forall X, Y \in T\mathcal{M}_n, \quad (3.146)$$

vagy másképp

$$-\frac{\partial^2}{\partial s \partial t} H_g(D + tX, D + sY) \Big|_{t=s=0} = \text{Tr} \left( X (R_{n,D}^{\frac{1}{2}}(\phi(g))(L_{n,D} R_{n,D}^{-1}) R_{n,D}^{\frac{1}{2}})^{-1} (Y) \right). \quad (3.147)$$

Továbbá  $g \in \mathcal{K}_{[0,\infty]}^{(n)}$  esetén  $\phi(g) \in \mathcal{F}_{[0,\infty]}^{(S,n)}$  teljesül.

Ebből látható, hogy minden  $g$ -relatív entrópia ugyan azt a metrikát generálja, mint a  $g$  szimmetrizáltjából származó relatív entrópia.

**3.27. Tétel.** Az

$$\epsilon : \mathcal{F}_{[0,\infty]}^{(S)} \rightarrow \mathcal{K}_{[0,\infty]}^{(S)} \quad f(x) \mapsto \epsilon(f)(x) = \frac{(x-1)^2}{2f(x)} \quad (3.148)$$

leképezés jól értelmezett, valamint  $K^{(n),f} = K^{\epsilon(f),(n)}$  teljesül, azaz minden  $D \in \mathcal{M}_n^+$  állapotra és  $X, Y \in T_D \mathcal{M}_n^+$  érintővektorra

$$-\frac{\partial^2}{\partial s \partial t} H_{\epsilon(f)}(D + tX, D + sY) \Big|_{t=s=0} = \text{Tr} \left( X (R_{n,D}^{\frac{1}{2}} f(L_{n,D} R_{n,D}^{-1}) R_{n,D}^{\frac{1}{2}})^{-1} (Y) \right). \quad (3.149)$$

Az eddigieket csoportosítva kapjuk az alábbi tételt.

**3.28. Tétel.** Kölcsönösen egyértelmű megfeleltetések léteznek az alábbi halmazok között.

1. Monoton metrikák.

2. Az  $\mathcal{F}_{[0,\infty]}^{(S)}$ -beli függvények.

3. A  $\mathcal{K}_{[0,\infty]}^{(S)}$ -beli függvények.

Vizsgáljuk meg, hogy mely  $g$ -relatív entrópiák generálják a főbb monoton metrikákat.

**3.13. Példa.** *Bures-féle (vagy legkisebb monoton) metrika:* Az  $f(x) = \frac{1+x}{2}$  függvényhez a (3.148) képlet alapján a

$$g(x) = \frac{(x-1)^2}{1+x} \quad (3.150)$$

$\mathcal{K}_{[0,\infty]}^{(S,n)}$ -beli függvény tartozik. Ezek szerint a Bures-metrikát a

$$H_{\text{SM}} : \mathcal{M}_n^+ \times \mathcal{M}_n^+ \rightarrow \mathbb{R} \quad (D_1, D_2) \mapsto H_{\text{SM}}(D_1, D_2) \quad (3.151)$$

$$H_{\text{SM}}(D_1, D_2) = \text{Tr}(D_1 - D_2)(L_{D_2} + R_{D_1})^{-1}(D_1 - D_2)$$

relatív entrópiá generálja. Ezt *Bures-féle relatív entrópiának* nevezzük.  $\spadesuit$

**3.14. Példa.** *Legnagyobb monoton metrika:* Az  $f(x) = \frac{2x}{1+x}$  függvényhez a (3.148) képlet alapján a

$$g(x) = (x-1)^2 \frac{1+x}{4x} \quad (3.152)$$

$\mathcal{K}_{[0,\infty]}^{(S,n)}$ -beli függvény tartozik. Azonban bevezetve a  $g_1(x) = \frac{(x-1)^2}{2} \mathcal{K}_{[0,\infty]}^{(n)}$ -beli függvényt, igazolható, hogy  $g(x) \approx g_1(x)$  teljesül. A 3.25. tétel alapján a  $g_1(x)$ -relatív entrópiá ugyan azt a monoton metrikát generálja mint a  $g$ -ből származtatott relatív entrópiá. Adott  $D_1, D_2 \in \mathcal{M}_n^+$  állapotokra  $H_{g_1}(D_1, D_2)$  értékét az alábbi módon határozhatjuk meg.

$$\begin{aligned} H_{g_1}(D_1, D_2) &= \frac{1}{2} \text{Tr} D_1^{1/2} (L_{D_2} R_{D_1}^{-1} - 1)^2 (D_1^{1/2}) = & (3.153) \\ &= \frac{1}{2} \text{Tr} D_1^{1/2} (D_2^2 D_1^{1/2} D_1^{-2} - 2D_2 D_1^{1/2} D_1^{-1} + D_1^{1/2}) = \\ &= \frac{1}{2} \text{Tr} D_2^2 D_1^{-1} - 2D_2 + D_1 = \\ &= \frac{1}{2} \text{Tr}(D_1 - D_2) D_1^{-1} (D_1 - D_2) \end{aligned}$$

A  $H_{g_1}$  mennyiség kétszerese  $H_{(x-1)^2}$ , ezt gyakran a *kvadratus relatív entrópiának* nevezik.  $\spadesuit$

**3.15. Példa.** *Kubo–Mori-metrika:* Az  $f(x) = \frac{x-1}{\log x}$  függvényhez a (3.148) képlet alapján a

$$g(x) = \frac{x-1}{2} \log x \quad (3.154)$$

$\mathcal{K}_{]0,\infty]}^{(S,n)}$ -beli függvény tartozik. A  $g_1(x) = -\log x$   $\mathcal{K}_{]0,\infty]}^{(n)}$ -beli függvényre  $g(x) \approx g_1(x)$  teljesül. A (3.25.) tétel alapján a  $g_1(x)$ -relatív entrópia ugyanazt a monoton metrikát generálja, mint a  $g$ -ből származtatott relatív entrópia. A

$$\log xy^{-1} = \int_0^\infty \frac{1}{y+t} - \frac{1}{x+t} dt \quad (3.155)$$

integrálreprezentációt használva az  $L_{D_2}$  és  $R_{D_1}^{-1}$  kommutáló mennyiségekre

$$\begin{aligned} H_{g_1}(D_1, D_2) &= \text{Tr } D_1^{1/2} \int_0^\infty (R_{D_1} + t)^{-1} (D_1^{1/2}) - (L_{D_2} + t)^{-1} (D_1^{1/2}) dt = \\ &= \text{Tr } D_1^{1/2} \int_0^\infty (D_1 + t)^{-1} D_1^{1/2} - D_1^{1/2} (D_2 + t)^{-1} dt = \\ &= \text{Tr } D_1 \int_0^\infty -((1+t)^{-1} - (D_1 + t)^{-1}) + ((1+t)^{-1} - (D_2 + t)^{-1}) dt = \\ &= \text{Tr } D_1 (\log D_1 - \log D_2) \end{aligned} \quad (3.156)$$

adódik. (A számolás folyamán a (3.155) egyenletet, a nyomképzés ciklikusságát, és a logaritmus (3.107) integrálreprezentációját használtuk fel.) Ezt a relatív entrópiát jelölje  $H_{\log}$ , melyet gyakran *Umegaki-féle relatív entrópiának* neveznek.  $\spadesuit$

**3.16. Példa.** *Wigner–Yanase–Dyson-metrika:* Az  $f_{\text{WYD}}$  függvényhez a (3.148) képlet alapján a

$$g(x) = \begin{cases} \frac{4}{1-\alpha^2} \left(1 - x^{\frac{1+\alpha}{2}}\right) & \text{ha } \alpha \neq \pm 1 \\ x \log x & \text{ha } \alpha = 1 \\ -\log x & \text{ha } \alpha = -1 \end{cases} \quad (3.157)$$

$\mathcal{K}_{]0,\infty]}^{(S,n)}$ -beli függvény tartozik.  $\spadesuit$

A  $g$ -relatív entrópiák közül bizonyos értelemben a Bures-féle a legkisebb és a kvadratikus a legnagyobb. Ezt fogalmazza meg pontosan az alábbi tétel [64].

**3.29. Tétel.** *Legyen  $g \in \mathcal{K}_{]0,\infty]}^{(n)}$  tetszőleges. Ekkor minden  $D_1, D_2 \in \mathcal{M}_n^+$  állapotra*

$$H_{(x-1)^2}^{(szim)}(D_1, D_2) \geq H_g(D_1, D_2) \geq H_{SM}(D_1, D_2) \quad (3.158)$$

*teljesül.*

A 2.7. pontban értelmeztük az eloszlások geometriai távolságát a 2.22. definícióval, mint adott metrika esetén a két eloszlást összekötő legrövidebb geodetikus hosszát. Ennek a nemkommutatív megfelelője az alábbi definíció.

**3.20. Definíció.** Legyen  $g \in \mathcal{G}_{[0,\infty]}^n$ . Az  $(\mathcal{M}_n^+, K^{g,(n)})$  Riemann-sokaságon értelmezett

$$d_g : \mathcal{M}_n^+ \times \mathcal{M}_n^+ \rightarrow \mathbb{R} \quad (D_1, D_2) \mapsto d_g(D_1, D_2) \quad (3.159)$$

$$d_g(D_1, D_2) = \inf \left\{ \int_0^1 \sqrt{K_{\gamma(t)}^{g,(n)}(\dot{\gamma}(t), \dot{\gamma}(t))} \, dt \mid \gamma : [0, 1] \rightarrow \mathcal{M}_n^+ \text{ sima függvény,} \right. \\ \left. \gamma(0) = D_1, \gamma(1) = D_2 \right\}$$

függvényt *g-geodetikus távolságnak* nevezzük.

**3.30. Tétel.** [64] Minden  $g \in \mathcal{G}_{[0,\infty]}^{(n)}$  függvényre  $D_g^2$ , a *g-geodetikus távolság négyzete*, monoton, differenciálható, szimmetrikus relatív entropikus távolság. Továbbá  $D_g$  teljesíti a háromszögegyenlőtleniséget, azaz

$$d_g(D_1, D_3) \leq d_g(D_1, D_2) + d_g(D_2, D_3) \quad \forall D_1, D_2, D_3 \in \mathcal{M}_n^+ \quad (3.160)$$

teljesül.

Nagyon kevés esetben ismert a *g-geodetikus távolságra* explicit képlet. Ezek közül az egyik a Bures-metrika esete, melynek főbb differenciálgeometriai jellemzői megtalálhatók Uhlmann [26, 104, 105, 106] munkáiban. A  $D_1$  és  $D_2$  állapotokat összekötő geodetikus hosszára a Bures-metrika esetén

$$d_{\text{Bures}}(D_1, D_2) = \sqrt{2 \left( 1 - \text{Tr}(D_1^{1/2} D_2 D_1^{1/2})^{1/2} \right)} \quad (3.161)$$

teljesül.

Az  $\alpha$ -konnexióhoz tartozó geodetikus távolságra Jencova vezetett le összefüggéseket [55].

A 3.29. tételhez hasonló teljesül a *g-geodetikus távolságra*.

**3.31. Tétel.** [64] Minden  $g \in \mathcal{G}_{[0,\infty]}^{(n)}$  függvényre és  $D_1, D_2 \in \mathcal{M}_n^+$  állapotokra

$$d_{(x-1)^2}^{\text{szim}}(D_1, D_2) \geq d_g(D_1, D_2) \geq 4d_{\text{Bures}}(D_1, D_2) \quad (3.162)$$

teljesül.



## 4. Az állapotér geometriája

A klasszikus statisztika geometriájáról szóló második fejezetben részletesen elemeztük a diszkrét eloszlás geometriáját. A harmadik fejezetben a diszkrét eloszlást általánosítottuk. A jelen fejezet célja az általánosított diszkrét eloszlás, a kvantummechanikai állapotér geometriájának a vizsgálata.

Az első részben a valós illetve komplex állapotér skalárgömbületét határozzuk meg. Kiderül, hogy ehhez nagy segítséget nyújt a normális eloszlások második fejezetben meghatározott geometriája, a skalárgömbületének a kiszámítása. Az állapotér skalárgömbületére kapott eredményből látjuk, hogy a klasszikus esettel ellentétben a skalárgömbület nem minden monoton metrika esetén állandó.

A második részben Petz sejtését vizsgáljuk meg közelebbről, mely szerint kevertebb állapotban nagyobb a skalárgömbület, ha az állapotot a Kubo–Mori-metrikával látjuk el. A sejtés alapja az, hogy a skalárgömbület összefüggésbe hozható egy állapot statisztikai megkülönböztethetőségével. A sejtés bizonyításának a nehézsége a skalárgömbületre kapott kifejezés bonyolultságában rejlik. Megmutatjuk, hogy a sejtés ekvivalens egy meglehetősen összetett egyenlőtlenséggel, melyet öt egyszerűbb tagra bontunk fel. Ezen tagok közül háromra bizonyítjuk az egyenlőtlenséget, a másik két esetben pedig még tovább bontjuk az egyenlőtlenségeket és bemutatjuk az azok bizonyításában eddig elért eredményeket. A rész végén megmutatjuk, hogy ha a sejtés igaz a komplex állapotok terén, akkor teljesül a valós állapotéren is.

A harmadik részben az  $\mathcal{M}_2^+$  állapotér skalárgömbületét elemezzük. Megadunk egy egyszerűbb kifejezést a skalárgömbület kiszámítására, melyet alkalmazunk is a főbb monoton metrikák esetére. Ezen vizsgált esetekben a skalárgömbületnek a legkevertebb állapotban globális maximuma van. Ellenpéldák segítségével megmutatjuk, hogy ez azonban nem igaz minden monoton metrikára. Végül megadjuk monoton metrikák egy újabb családját, mely folytonos utat képez a legnagyobb illetve legkisebb metrika között.

A negyedik részben az  $\mathcal{M}_3^+$  illetve  $\mathcal{M}_4^+$  állapotér skalárgömbületével kapcsolatos numerikus szimulációkat mutatjuk be. Ezekből kiderül többek között, hogy a fontosabb monoton metrikák között több olyan is van, melyhez tartozó skalárgömbület nem lesz monoton a majorizációra nézve az  $\mathcal{M}_3^+$  esetben.

Végül az ötödik részben az  $\mathcal{M}_2^+$  tér térfogatát határozzuk meg, különböző monoton metrikák esetén, felírjuk a geodetikussok egyenletét egy megoldás bemutatásával. A legkevertebb állapot körüli gömb térfogatának megadjuk a sugár szerinti Taylor-sorfejtését tetszőleges monoton metrikára, majd megvizsgáljuk, hogyan változik a sorfejtés a főbb monoton metrikák esetében, ha a gömb középpontja nem a legkevertebb állapot.

### 4.1. Az állapotér skalárgömbülete

A monoton metrikák segítségével Riemann-sokasággá lehet tenni az  $\mathcal{M}_n^+$  állapotteret. Az  $(\mathcal{M}_n^+, K^{(n)})$  Riemann-tér differenciálgeometriai jellemzőinek a vizsgálata az 1990-es években kezdődött. A sokaság skalárgömbületét először Petz [81] cikke említi, ott az  $(\mathcal{M}_2^+, K_{\text{KM}}^2)$  Riemann-sokaság esetére meg is lett határozva. A gömbülete vonatkozó következő eredményt Petz és Sudár publikálta 1996-ban [88], ahol az  $\mathcal{M}_2^+$  tér metszetgömbületeit határozták meg. Az  $(\mathcal{M}_n^+, K^{(n)})$  tér skalárgömbületét a Kubo–Morimetrika mellett valós állapotok esetén Michor, Petz és Andai [73], komplex állapotok esetén pedig Dittmann [25] számolta ki. A jelen részben a valós állapotokra elvégzett számításokat terjesztjük ki tetszőleges monoton metrikára és komplex állapotokra is.

Az  $(\mathcal{M}_n^+, K^{(n),f})$  Riemann-tér skalárgömbületének a kiszámítása a normális eloszlások geometriájának a vizsgálatánál, illetve a részsokaságok skalárgömbületénél leírt számításokon alapul.

A 2.4. példát követve először bevezetjük az  $(\tilde{\mathcal{M}}_n^+, \tilde{K}^{(n),f})$  Riemann-geometriát, melynek egy-kodimenziós részsokasága lesz az  $(\mathcal{M}_n^+, K^{(n),f})$  tér. Majd a 2.22. tétel alapján határozzuk meg az állapotér skalárgömbületét.

Jelölje  $\tilde{\mathcal{M}}_n^+$  az  $n \times n$ -es önadjungált, pozitív definit valós vagy komplex mátrixok halmazát. Az  $\tilde{\mathcal{M}}_n^+$  halmaz részhalmaza az  $\mathbb{R}^{(n+2)(n-1)/2}$  vagy  $\mathbb{R}^{n^2-1}$  térnek, attól függően, hogy valós illetve komplex elemeket tartalmaz. Így természetes módon ellátható differenciálható struktúrával. Ha  $\tilde{D} \in \tilde{\mathcal{M}}_n^+$ , akkor a  $T_{\tilde{D}}\tilde{\mathcal{M}}_n^+$  érintőtér azonosítható az  $n \times n$ -es, önadjungált valós vagy komplex mátrixok halmazával.

Adott  $f \in \mathcal{F}_{[0,\infty]}^{(S,n)}$  függvényre értelmezzük a

$$\tilde{K}^{(n),f} : \tilde{\mathcal{M}}_n^+ \rightarrow \text{Lin}(T\tilde{\mathcal{M}}_n \times T\tilde{\mathcal{M}}_n, \mathbb{R}) \quad \tilde{D} \mapsto \left( (\tilde{X}, \tilde{Y}) \mapsto \tilde{K}_{\tilde{D}}^{(n),f}(\tilde{X}, \tilde{Y}) \right) \quad (4.1)$$

leképezést minden  $\tilde{D} \in \tilde{\mathcal{M}}_n^+$  esetén minden  $\tilde{X}, \tilde{Y} \in T_{\tilde{D}}\tilde{\mathcal{M}}_n^+$  érintővektorra a

$$\tilde{K}_{\tilde{D}}^{(n),f}(\tilde{X}, \tilde{Y}) = \text{Tr} \left( \tilde{X} (R_{n,\tilde{D}}^{\frac{1}{2}} f(L_{n,\tilde{D}} R_{n,\tilde{D}}^{-1}) R_{n,\tilde{D}}^{\frac{1}{2}})^{-1} (\tilde{Y}) \right) \quad (4.2)$$

képlettel. Ekkor az  $(\tilde{\mathcal{M}}_n^+, \tilde{K}^{(n),f})$  pár Riemann-geometria.

Az

$$i : \mathcal{M}_n^+ \rightarrow \tilde{\mathcal{M}}_n^+ \quad D \mapsto D \quad (4.3)$$

identikus leképezés beágyazás, továbbá az  $\mathcal{M}_n^+$  tér egy kodimenziós részsokasága  $\tilde{\mathcal{M}}_n^+$ -nek. Továbbá a  $\tilde{K}^{(n),f}$  metrika  $i$ -vel való visszahúzottja megegyezik a  $K^{(n),f}$  metrikával.



Tekintsük az

$$n : \mathcal{M}_n^+ \rightarrow T\tilde{\mathcal{M}}_n \quad D \mapsto n(D) = D \quad (4.4)$$

leképezést. Ez az  $\mathcal{M}_n^+$  sokaság normálvektormezője, ugyanis minden  $D \in \mathcal{M}_n^+$  pont és  $X \in T_D\mathcal{M}_n^+$  érintővektor esetén

$$\tilde{K}(X, n(D)) = \text{Tr} \left( X (R_{n,D}^{\frac{1}{2}} f(L_{n,D} R_{n,D}^{-1}) R_{n,D}^{\frac{1}{2}})^{-1}(D) \right) = \text{Tr} X f(1) = \text{Tr} X = 0, \quad (4.5)$$

$$\tilde{K}(n(D), n(D)) = \text{Tr} \left( D (R_{n,D}^{\frac{1}{2}} f(L_{n,D} R_{n,D}^{-1}) R_{n,D}^{\frac{1}{2}})^{-1}(D) \right) = \text{Tr} D f(1) = \text{Tr} D = 1$$

teljesül. A számolás során kihasználtuk, hogy a  $D$  operátor különböző hatványai felcserélhetőek egymással, hogy az  $f$  függvény normált, és hogy a  $T_D\mathcal{M}_n^+$  érintőtér elemeinek nulla a nyoma.

A 2.22. tétel feltételei teljesülnek az  $\tilde{\mathcal{M}}_n^+$ ,  $\mathcal{M}_n^+$  sokaságokra, ezért a (2.102) képlet segítségével fogjuk az  $(\mathcal{M}_n^+, K^{(n),f})$  Riemann-sokaság skalárgörbületét meghatározni.

A (2.102) képletben szereplő Riemann-féle görbületi tenzor és az  $S$  leképezés meghatározásához a 3.4. részben bemutatott Riesz–Dunford-féle operátorkalkulust fogjuk használni.

A (2.124) képletnek megfelelően vezessük be a

$$\tilde{G}_f : \tilde{\mathcal{M}}_n^+ \rightarrow \text{Lin}(T\tilde{\mathcal{M}}_n, T\tilde{\mathcal{M}}_n) \quad \tilde{D} \mapsto (\tilde{X} \mapsto \tilde{G}_f(\tilde{D})(\tilde{X})) \quad (4.6)$$

$$\tilde{G}_f(\tilde{D})(\tilde{X}) = \text{Tr} \frac{1}{(2\pi i)^2} \oint \oint c_f(\xi, \eta) \tilde{X}(\xi - \tilde{D})^{-1} \tilde{Y}(\eta - \tilde{D})^{-1} d\xi d\eta$$

leképezést, ahol  $c_f$  jelöli az  $f$  függvény által meghatározott Cencov–Morozova-féle függvényt. Ekkor minden  $\tilde{D} \in \tilde{\mathcal{M}}_n^+$  pontban, minden  $\tilde{X}, \tilde{Y} \in T_{\tilde{D}}\tilde{\mathcal{M}}_n^+$  érintővektorra

$$\tilde{K}_{\tilde{D}}^{(n),f}(\tilde{X}, \tilde{Y}) = \text{Tr}(\tilde{X}\tilde{G}(\tilde{D})(\tilde{Y})) \quad (4.7)$$

teljesül a (2.123) és a (2.125) képletnek megfelelően.

Ezen a ponton látható, hogy a „nem normált” állapotér  $(\tilde{\mathcal{M}}_n^+, \tilde{K}^{(n),f})$  differenciálgeometriája nagyfokú hasonlóságot mutat a többdimenziós normális eloszlás geometriájához. Az ott elvégzett számításokat a jelen esetben is alkalmazhatjuk az  $\int_T = \frac{1}{(2\pi i)^2} \oint \oint$  és a  $d\mu(t) = c(\xi, \eta) d\xi d\eta$  formális helyettesítés után.

A (2.127) képlettel értelmezett kétváltozós  $m$  függvény, ekkor

$$m : \mathbb{R}^+ \times \mathbb{R}^+ \rightarrow \mathbb{R} \quad (x, y) \mapsto m(x, y) = \frac{1}{(2\pi i)^2} \oint \oint c_f(\xi, \eta) \frac{1}{\xi - x} \frac{1}{\eta - y} d\xi d\eta, \quad (4.8)$$

mely szerint  $m(x, y) = c_f(x, y)$  teljesül. A  $c_f$  függvény szimmetrikussága miatt  $m_f(x, y) = m_f(y, x)$  teljesül, továbbá az  $f$  függvény normáltsága miatt  $m_f(x, x) = \frac{1}{x}$ .

A 2.24. tétel az  $\tilde{\mathcal{M}}_n^+$  sokaság esetén az alábbi formában érvényes.

**4.1. Tétel.** *Tekintsük az  $(\tilde{\mathcal{M}}_n^+, \tilde{K}^{(n),f})$  sokaságot, ahol  $\tilde{K}^{(n),f}$  az  $f \in \mathcal{F}_{[0,\infty]}^{(S,n)}$  függvény által generált metrika. Legyen a  $\tilde{D} \in \tilde{\mathcal{M}}_n^+$  mátrix  $\tilde{D} = \sum_{k=1}^n \lambda_k E_{kk}$  alakú.*

$$\begin{aligned} \text{Ha } 1 \leq i < j \leq n, \quad 1 \leq k < l \leq n, \text{ akkor:} & \begin{cases} G_D(H_{ij}, H_{kl}) = \delta_{ik}\delta_{jl}2m(\lambda_i, \lambda_j) \\ G_D(F_{ij}, F_{kl}) = \delta_{ik}\delta_{jl}2m(\lambda_i, \lambda_j) \\ G_D(H_{ij}, F_{kl}) = 0, \end{cases} \\ \text{ha } 1 \leq i < j \leq n, \quad 1 \leq k \leq n, \text{ akkor:} & G_D(H_{ij}, F_{kk}) = G(F_{ij}, F_{kk}) = 0, \quad (4.9) \\ \text{ha } 1 \leq i \leq n, \quad 1 \leq k \leq n, \text{ akkor:} & G_D(F_{ii}, F_{kk}) = \delta_{ik}4m(\lambda_i, \lambda_i) \end{aligned}$$

teljesül.

A (2.143) formulához hasonlóan a  $\tilde{G}_f$  függvény deriváltja

$$d\tilde{G} : \tilde{\mathcal{M}}_n^+ \rightarrow \text{Lin}(T\tilde{\mathcal{M}}_n, \text{Lin}(T\tilde{\mathcal{M}}_n, T\tilde{\mathcal{M}}_n)) \quad \tilde{D} \mapsto \left( \tilde{Y} \mapsto (\tilde{X} \mapsto d\tilde{G}(\tilde{D})(\tilde{Y})(\tilde{X})) \right) \quad (4.10)$$

$$d\tilde{G}(\tilde{D})(\tilde{Y})(\tilde{X}) = -\frac{1}{(2\pi i)^2} \oint \oint (\xi - \tilde{D})^{-1} \tilde{X} (\xi - \tilde{D})^{-1} \tilde{Y} (\eta - \tilde{D})^{-1} \\ + (\xi - \tilde{D})^{-1} \tilde{Y} (\eta - \tilde{D})^{-1} \tilde{X} (\eta - \tilde{D})^{-1} d\xi d\eta.$$

lesz. Ezek alapján a (2.145) képlettel értelmezett háromváltozós  $m$  függvény itt az

$$m : \mathbb{R}^+ \times \mathbb{R}^+ \times \mathbb{R}^+ \rightarrow \mathbb{R} \quad (x, y, z) \mapsto m(x, y, z) \quad (4.11)$$

$$m(x, y, z) = -\frac{1}{(2\pi i)^2} \oint \oint c_f(\xi, \eta) \frac{1}{(\xi - a)(\eta - b)(\xi - c)} d\xi d\eta$$

alakot ölti. Ezek alapján a (2.148) egyenlet továbbra is érvényes.

Az előző fejezetekben alkalmazott jelölésmóddhoz hasonlóan, amennyiben a  $\tilde{D} \in \tilde{\mathcal{M}}_n^+$  mátrix diagonális, és a főátlóban a  $\lambda_1, \dots, \lambda_n$  számok állnak, akkor az  $1 \leq i, j, k \leq n$  indexek esetén az

$$m_{ij} = m(\lambda_i, \lambda_j) \quad m_{ijk} = m(\lambda_i, \lambda_j, \lambda_k) \quad (4.12)$$

rövidítést fogjuk használni.

Komplex állapotter esetén a (2.149) egyenlet, csak az  $(F_{ij})_{1 \leq i \leq j \leq n}$  érintővektorokon adja meg a  $d\tilde{G}$  leképezés értékét. A komplex elemeket tartalmazó  $(H_{ij})_{1 \leq i < j \leq n}$  érintővektorok esetén az alábbi formulák adják meg a  $d\tilde{G}$  leképezés értékét.

$$d\tilde{G}(\tilde{D})(H_{ij})(H_{kl}) = (F_{il}m_{ijl}\delta_{jk} + F_{jk}m_{ijk}\delta_{il} - F_{ik}m_{ijk}\delta_{jl} - F_{jl}m_{ijl}\delta_{ik}) \quad (4.13)$$

$$d\tilde{G}(\tilde{D})(F_{ij})(H_{kl}) = -(H_{il}m_{ijl}\delta_{jk} + H_{kj}m_{ijk}\delta_{il} - H_{ik}m_{ijk}\delta_{jl} - H_{lj}m_{ijl}\delta_{ik})$$

A  $c_f$  függvényre minden pozitív  $x, y, t \in \mathbb{R}$  paraméter esetén

$$c_f(x, y) = tc_f(tx, ty) \quad (4.14)$$

teljesül. Ennek az egyenletnek a  $t$  szerinti parciális deriváltjában a  $t = 1$  értéknél az

$$c_f(x, y) + x\partial_1 c_f(x, y) + y\partial_2 c_f(x, y) = 0 \quad (4.15)$$

adódik. Az egyenlet másik alakja az alábbi

$$\frac{1}{(2\pi i)^2} \oint \oint c_f(\xi, \eta) \frac{1}{(\xi - x)(\eta - y)} \left(1 + \frac{x}{\xi - x} + \frac{y}{\eta - y}\right) = 0. \quad (4.16)$$

Felhasználva az  $m$  függvények tulajdonságait ebből az

$$\frac{m(x, x, y)}{m(x, x)} + \frac{m(x, y, y)}{m(y, y)} = m(x, y) \quad (4.17)$$

azonosságot kapjuk.

A 2.6. részben a Levi–Civita-féle kovariáns deriválásra kapott (2.157) kifejezés továbbra is érvényes, vagyis

$$\tilde{\Gamma} : \tilde{\mathcal{M}}_n^+ \rightarrow \text{Lin}(T\tilde{\mathcal{M}}_n \times T\tilde{\mathcal{M}}_n, T\tilde{\mathcal{M}}_n) \quad \tilde{D} \mapsto ((\tilde{X}, \tilde{Y}) \mapsto \tilde{\Gamma}(\tilde{D})(\tilde{X}, \tilde{Y})) \quad (4.18)$$

$$\tilde{\Gamma}(\tilde{D})(\tilde{X}, \tilde{Y}) = \frac{1}{2} \tilde{G}^{(-1)}(\tilde{D}) \left( d\tilde{G}(\tilde{D})(\tilde{X})(\tilde{Y}) \right).$$

Ezek alapján meghatározhatjuk a részsokaság skalárgörbületének kiszámításánál felhasználandó (2.102) képletben szereplő  $S$  leképezést, melyet a (2.100) képlet definiál. Legyen  $D \in \mathcal{M}_n^+$  és  $X, Y \in T_D \mathcal{M}_n^+$ , ekkor

$$\begin{aligned} S(X, Y) &= -\tilde{K}_D^{(n),f}(\tilde{\Gamma}(D)(X)(D), Y) = -\tilde{K}_D^{(n),f} \left( \frac{1}{2} \tilde{G}^{(-1)}(D) \left( d\tilde{G}(D)(X)(D) \right), Y \right) = \\ &= -\frac{1}{2} \text{Tr} \left( d\tilde{G}(D)(X)(D)Y \right) = \frac{1}{2} \tilde{K}_D^{(n),f}(X, Y). \end{aligned}$$

Ebből egyúttal az

$$\tilde{\Gamma}(D)(X)(D) = -\frac{1}{2}X \quad \forall D \in \mathcal{M}_n^+ \quad \forall X \in T_D\mathcal{M}_n^+ \quad (4.19)$$

egyenlet is következik.

Az  $(\tilde{\mathcal{M}}_n^+, \tilde{K}^{(n),f})$  Riemann-tér görbületi tenzorát pedig a (2.166) kifejezés adja meg.

A fenti képletek felhasználásával a 2.22. tételnek megfelelően könnyen ki tudjuk számolni az  $(\mathcal{M}_n^+, K^{(n),f})$  tér skalárgörbületét.

Adott  $D \in \mathcal{M}_n^+$  állapot esetén legyen  $(B_t)_{t \in I}$  a  $T_D\mathcal{M}_n^+$  érintőtér ortonormált bázisa a  $K_D^{(n),f}$  skalárszorozásra nézve. Ekkor a  $D \in T\tilde{\mathcal{M}}_n$  vektor, mivel normálvektor, merőleges minden  $B_t$  vektorra és egységnyi hosszú, vagyis a  $B_0 = D$  és  $I_0 = I \cup \{0\}$  jelöléssel élve a  $(B_t)_{t \in I_0}$  vektorrendszer ortonormált bázis a  $T_D\tilde{\mathcal{M}}_n^+$  térben.

A 2.22. tételben szereplő első összegzést az alábbi formába írhatjuk át.

$$\begin{aligned} & \sum_{t,s \in I} \tilde{K}_D^{(n),f}(\tilde{R}_D(A_t, A_s)A_s, A_t) = \quad (4.20) \\ & = \sum_{t,s \in I_0} \tilde{K}_D^{(n),f}(\tilde{R}_D(A_t, A_s)A_s, A_t) - \sum_{t \in I} \tilde{K}_D^{(n),f}(\tilde{R}_D(A_t, D)D, A_t) - \\ & \quad - \sum_{s \in I} \tilde{K}_D^{(n),f}(\tilde{R}_D(D, A_s)A_s, D) - \tilde{K}_D^{(n),f}(\tilde{R}_D(D, D)D, D) = \\ & = \tilde{\text{Scal}}(D) - \sum_{t \in I} \left( \tilde{K}_D^{(n),f}(\tilde{R}_D(A_t, D)D, A_t) + \tilde{K}_D^{(n),f}(\tilde{R}_D(D, A_t)A_t, D) \right) - \\ & \quad - \tilde{K}_D^{(n),f}(\tilde{R}_D(D, D)D, D) \quad (4.21) \end{aligned}$$

Igazoljuk, hogy ez a kifejezés megegyezik az  $\tilde{\mathcal{M}}_n^+$  sokaság  $D$  pontbeli  $\tilde{\text{Scal}}(D)$  skalárgörbületével. Az  $\tilde{R}$  Riemann-féle görbületi tenzor meghatározásához a (2.166) képletet használjuk. A (4.21) egyenlet első összeadandójáról az alábbi módon igazolhatjuk, hogy nulla.

$$\tilde{K}_D^{(n),f}(\tilde{R}_D(A_t, D)D, A_t) = \quad (4.22)$$

$$\begin{aligned}
&= \tilde{K}_D^{(n),f} \left( \frac{1}{4} \tilde{G}^{(-1)}(D) \left( d\tilde{G}(D)(D) \left( \tilde{G}^{(-1)}(D) (d\tilde{G}(D)(A_t)(D)) \right) \right), A_t \right) - \\
&\quad - \tilde{K}_D^{(n),f} \left( \frac{1}{4} \tilde{G}^{(-1)}(D) \left( d\tilde{G}(D)(A_t) \left( \tilde{G}^{(-1)}(D) (d\tilde{G}(D)(D)(D)) \right) \right), A_t \right) = \\
&= \frac{1}{4} \operatorname{Tr} \left( d\tilde{G}(D)(D) \left( \tilde{G}^{(-1)}(D) (d\tilde{G}(D)(A_t)(D)) \right) A_t \right) - \\
&\quad - \frac{1}{4} \operatorname{Tr} \left( d\tilde{G}(D)(A_t) \left( \tilde{G}^{(-1)}(D) (d\tilde{G}(D)(D)(D)) \right) A_t \right) = \\
&= \frac{1}{4} \operatorname{Tr} \left( d\tilde{G}(D)(D) (2\tilde{\Gamma}(D)(A_t)(D)) A_t \right) - \frac{1}{4} \operatorname{Tr} \left( d\tilde{G}(D)(A_t) (2\tilde{\Gamma}(D)(D)(D)) A_t \right) = \\
&= \frac{1}{4} \operatorname{Tr} (d\tilde{G}(D)(D)(-A_t)A_t) - \frac{1}{4} \operatorname{Tr} (d\tilde{G}(D)(A_t)(-D)A_t) = 0
\end{aligned}$$

A (4.21) egyenlet második összeadandója szintén nulla lesz.

$$\begin{aligned}
&\tilde{K}_D^{(n),f} (\tilde{R}_D(D, A_t)A_t, D) = \tag{4.23} \\
&= \tilde{K}_D^{(n),f} \left( \frac{1}{4} \tilde{G}^{(-1)}(D) \left( d\tilde{G}(D)(A_t) \left( \tilde{G}^{(-1)}(D) (d\tilde{G}(D)(D)(A_t)) \right) \right), D \right) - \\
&\quad - \tilde{K}_D^{(n),f} \left( \frac{1}{4} \tilde{G}^{(-1)}(D) \left( d\tilde{G}(D)(D) \left( \tilde{G}^{(-1)}(D) (d\tilde{G}(D)(A_t)(A_t)) \right) \right), D \right) = \\
&= \frac{1}{4} \operatorname{Tr} \left( d\tilde{G}(D)(A_t) \left( \tilde{G}^{(-1)}(D) (d\tilde{G}(D)(D)(A_t)) \right) D \right) - \\
&\quad - \frac{1}{4} \operatorname{Tr} \left( d\tilde{G}(D)(D) \left( \tilde{G}^{(-1)}(D) (d\tilde{G}(D)(A_t)(A_t)) \right) D \right) = \\
&= \frac{1}{4} \operatorname{Tr} \left( d\tilde{G}(D)(A_t) (2\tilde{\Gamma}(D)(A_t)(D)) D \right) - \\
&\quad - \frac{1}{4} \operatorname{Tr} \left( 2\tilde{G}(D) \left( \tilde{\Gamma}(D) \left( \tilde{G}^{(-1)}(D) (d\tilde{G}(D)(A_t)(A_t)) \right) \right) (D) \right) = \\
&= \frac{1}{4} \operatorname{Tr} (d\tilde{G}(D)(D)(-A_t)A_t D) + \frac{1}{4} \operatorname{Tr} \left( \tilde{G}(D) \left( \tilde{G}^{(-1)}(D) (d\tilde{G}(D)(A_t)(A_t)) \right) D \right) =
\end{aligned}$$

$$= -\frac{1}{4} \operatorname{Tr}(d\tilde{G}(D)(D)(A_t)A_tD) + \frac{1}{4} \operatorname{Tr}(d\tilde{G}(D)(D)(A_t)A_tD) = 0$$

A (4.21) egyenlet negyedik tagja pedig triviálisan nulla.

A (2.102) egyenletben szereplő második illetve harmadik összeadandót külön határozzuk meg valós illetve komplex állapottérre. Egy  $D \in \mathcal{M}_n^+$  ponthoz tartozó  $T_D\mathcal{M}_n^+$  érintőtér valós esetben  $d_{\mathbb{R}}(n) = \frac{(n-1)(n+2)}{2}$ , komplex esetben pedig  $d_{\mathbb{C}}(n) = n^2 - 1$  dimenziós. Legyen  $(A_i)_{i=1,\dots,d}$  ortonormált bázis a  $\tilde{K}_D^{(n),f}$  skalárszorzásra nézve. Ekkor a (2.102) képletben szereplő második-, harmadik összegzés az alábbi.

$$\sum_{t,s=1}^d \left( S(A_s, A_s)S(A_t, A_t) - S(A_t, A_s)S(A_s, A_t) \right) = \sum_{t,s=1}^d \frac{1}{4} - \frac{1}{4} \delta_{t,s} = \frac{1}{4}(d^2 - d) \quad (4.24)$$

A (2.102) képlet ezek alapján az

$$\operatorname{Scal}(D) = \tilde{\operatorname{Scal}}(D) + \frac{d(d-1)}{4} \quad (4.25)$$

alakot ölti, ahol  $d = d_{\mathbb{R}}(n)$  vagy  $d = d_{\mathbb{C}}(n)$ , attól függően, hogy az állapottér valós vagy komplex.

Valós állapottér esetén a  $D \in \tilde{\mathcal{M}}_n^+$  pontban a  $(\tilde{\mathcal{M}}_n^+, \tilde{K}^{(n),f})$  Riemann-sokaság  $\tilde{\operatorname{Scal}}(D)$  skalárgörbületét a (2.227) képlet vagy az egyszerűsítések után kapott (2.229) kifejezés adja meg.

Komplex állapottérnél a helyzet annyival bonyolultabb, hogy a  $D \in \tilde{\mathcal{M}}_n^+$  állapot  $T_D\tilde{\mathcal{M}}_n^+$  érintőterében az  $\left( \frac{F_{ij}}{\sqrt{2}}, \frac{H_{ij}}{\sqrt{2}} \right)_{1 \leq i < j \leq n}$ ,  $\left( \frac{F_{ii}}{2} \right)_{1 \leq i \leq n}$  vektorrendszer lesz ortonormált bázis a  $\tilde{K}_D^{(n),f}$  skalárszorzásra nézve. A 2.6. részben bevezetett indexhalmazokat használjuk most is, vagyis  $I_1 = \{(i, j) \mid 1 \leq i < j \leq n\}$  és  $I_2 = \{(i, i) \mid 1 \leq i \leq n\}$ . Ha  $t = (i, j) \in I_1$ , akkor legyen  $A_t = \frac{F_{ij}}{\sqrt{2}}$ ,  $B_t = \frac{H_{ij}}{\sqrt{2}}$ ,  $t = (i, i) \in I_2$  esetén pedig  $A_t = \frac{F_{ii}}{2}$ . A  $\tilde{\operatorname{Scal}}(D)$  skalárgörbületet a (2.167) képlettel számoljuk ki. Az ott megjelenő összegzést az alábbi módon bontjuk fel.

$$\begin{aligned}
\tilde{\text{Scal}}(D) &= \sum_{t,s \in I} \left\langle \tilde{R}(D)(A_s, A_t) \left( \tilde{G}^{(-1)}(D)(A_t) \right), A_s \right\rangle + & (4.26) \\
&+ \sum_{\substack{t \in I_1 \\ s \in I_1}} \left\langle \tilde{R}(D)(A_s, B_t) \left( \tilde{G}^{(-1)}(D)(B_t) \right), A_s \right\rangle + \\
&+ \sum_{\substack{t \in I_2 \\ s \in I_1}} \left\langle \tilde{R}(D)(B_s, A_t) \left( \tilde{G}^{(-1)}(D)(A_t) \right), B_s \right\rangle + \\
&+ \sum_{\substack{t \in I_1 \\ s \in I_1}} \left\langle \tilde{R}(D)(B_s, A_t) \left( \tilde{G}^{(-1)}(D)(A_t) \right), B_s \right\rangle + \\
&+ \sum_{\substack{t \in I_1 \\ s \in I_2}} \left\langle \tilde{R}(D)(A_s, B_t) \left( \tilde{G}^{(-1)}(D)(B_t) \right), A_s \right\rangle + \\
&+ \sum_{t,s \in I_1} \left\langle \tilde{R}(D)(B_s, B_t) \left( \tilde{G}^{(-1)}(D)(B_t) \right), B_s \right\rangle
\end{aligned}$$

A felbontásban szereplő első tagot számoltuk ki több kisebb részletben a 2.6 részben. Most a többi tagot határozzuk meg.

A (4.26) képlet második összege:

$$\sum_{\substack{t \in I_1 \\ s \in I_1}} \left\langle \tilde{R}(D)(A_s, B_t) \left( \tilde{G}^{(-1)}(D)(B_t) \right), A_s \right\rangle = \quad (4.27)$$

$$= \sum_{\substack{1 \leq i < j \leq n \\ 1 \leq k < l \leq n}} \left\langle \tilde{R}(D) \left( \frac{F_{kl}}{\sqrt{2}}, \frac{H_{ij}}{\sqrt{2}} \right) \frac{1}{m_{ij}} \frac{H_{ij}}{\sqrt{2}}, \frac{F_{kl}}{\sqrt{2}} \right\rangle = \quad (4.28)$$

$$= \sum_{\substack{1 \leq i < j \leq n \\ 1 \leq k < l \leq n}} \frac{1}{4m_{ij}} \langle \tilde{R}(D)(F_{kl}, H_{ij}) H_{ij}, F_{kl} \rangle = \quad (4.29)$$

$$\begin{aligned}
&= \sum_{\substack{1 \leq i < j \leq n \\ 1 \leq k < l \leq n}} \frac{-1}{16m_{ij}} \left\langle \tilde{G}^{(-1)}(D) \left( d\tilde{G}(D)(F_{kl}) \left( \tilde{G}^{(-1)}(D) \left( d\tilde{G}(D)(H_{ij})(H_{ij}) \right) \right) \right), F_{kl} \right\rangle + \\
&\quad (4.30) \\
&+ \sum_{\substack{1 \leq i < j \leq n \\ 1 \leq k < l \leq n}} \frac{1}{16m_{ij}} \left\langle \tilde{G}^{(-1)}(D) \left( d\tilde{G}(D)(H_{ij}) \left( \tilde{G}^{(-1)}(D) \left( d\tilde{G}(D)(F_{kl})(H_{ij}) \right) \right) \right), F_{kl} \right\rangle .
\end{aligned}$$

A (4.30) képlet első tagja:

$$\sum_{\substack{1 \leq i < j \leq n \\ 1 \leq k < l \leq n}} \frac{-1}{16m_{ij}} \left\langle \tilde{G}^{(-1)}(D) \left( d\tilde{G}(D)(F_{kl}) \left( \tilde{G}^{(-1)}(D) (-F_{ii}m_{ij} - m_{ij}F_{jj}) \right) \right), F_{kl} \right\rangle . \quad (4.31)$$

Ez megegyezik a (2.173) képlettel, vagyis ezen összeg a (2.180) képlet alapján

$$= -\frac{1}{2} \left[ \sum_{1 \leq u < v < w \leq n} \left( \frac{m_{vvw}m_{uvv}}{m_{vu}m_{vv}m_{vw}} + \frac{m_{uuv}m_{uvv}}{m_{uu}m_{uv}m_{uw}} + \frac{m_{wvu}m_{wvv}}{m_{wu}m_{wv}m_{ww}} \right) + \quad (4.32)$$

$$+ \sum_{1 \leq k < l \leq n} \left( \frac{m_{kll}^2}{m_{kl}^2 m_{ll}} + \frac{m_{kkl}^2}{m_{kl}^2 m_{kk}} \right) \right] . \quad (4.33)$$

A (4.30) képlet második tagja:

$$\sum_{\substack{1 \leq i < j \leq n \\ 1 \leq k < l \leq n}} \frac{1}{16m_{ij}} \left\langle \tilde{G}^{(-1)}(D) \left( d\tilde{G}(D)(H_{ij}) \left( \tilde{G}^{(-1)}(D) \left( d\tilde{G}(D)(F_{kl})(H_{ij}) \right) \right) \right), F_{kl} \right\rangle = \quad (4.34)$$

$$\begin{aligned}
&= \sum_{\substack{1 \leq i < j \leq n \\ 1 \leq k < l \leq n}} -\frac{1}{16m_{ij}} \left\langle \tilde{G}^{(-1)}(D) \left( d\tilde{G}(D)(H_{ij}) \left( \tilde{G}^{(-1)}(D) (H_{il}m_{ijl}\delta_{jk} + H_{kj}m_{ijk}\delta_{il} + \right. \right. \right. \\
&\quad \left. \left. \left. + H_{ik}m_{ijk}\delta_{jl} + H_{lj}m_{ijl}\delta_{ik}) \right) \right), F_{kl} \right\rangle = \\
&\quad (4.35)
\end{aligned}$$



$$\begin{aligned}
&= \sum_{\substack{1 \leq i < j \leq n \\ 1 \leq k < l \leq n}} \frac{-1}{16m_{ij}} \left\langle \tilde{G}^{(-1)}(D) \left( d\tilde{G}(D)(H_{ij})(H_{il}) \frac{m_{ijl}}{m_{il}} \delta_{jk} + d\tilde{G}(D)(H_{ij})(H_{kj}) \frac{m_{ijk}}{m_{jk}} \delta_{il} + \right. \right. \\
&\quad \left. \left. + d\tilde{G}(D)(H_{ij})(H_{ik}) \frac{m_{ijk}}{m_{ik}} \delta_{jl} + d\tilde{G}(D)(H_{ij})(H_{lj}) \frac{m_{ijl}}{m_{jl}} \delta_{ik} \right), F_{kl} \right\rangle = \\
&= \sum_{\substack{1 \leq i < j \leq n \\ 1 \leq k < l \leq n}} -\frac{1}{16m_{ij}} \times \\
&\quad \times \left\langle \tilde{G}^{(-1)}(D) \left( \frac{m_{ijl}}{m_{il}} \delta_{jk} (-F_{ii}m_{iij} \delta_{jl} + F_{il}m_{ijl} \delta_{ji} + F_{ji}m_{iji} \delta_{il} - F_{jl}m_{ijl} \delta_{ii}) + \right. \right. \\
&\quad + \frac{m_{ijk}}{m_{jk}} \delta_{il} (F_{ij}m_{ijj} \delta_{jk} - F_{ik}m_{ijk} \delta_{jj} - F_{jj}m_{ijj} \delta_{ik} + F_{jk}m_{ijk} \delta_{ij}) + \\
&\quad + \frac{m_{ijk}}{m_{ik}} \delta_{jl} (-F_{ii}m_{iij} \delta_{jk} + F_{ik}m_{ijk} \delta_{ji} + F_{ji}m_{iji} \delta_{ik} - F_{jk}m_{ijk} \delta_{ii}) + \\
&\quad \left. \left. + \frac{m_{ijl}}{m_{jl}} \delta_{ik} (F_{ij}m_{ijj} \delta_{jl} - F_{il}m_{ijl} \delta_{jj} - F_{jj}m_{ijj} \delta_{il} + F_{jl}m_{ijl} \delta_{ij}) \right), F_{kl} \right\rangle = \\
&= -\frac{1}{8} \sum_{\substack{1 \leq i < j \leq n \\ 1 \leq k < l \leq n}} \frac{1}{m_{ij}} \left[ \frac{m_{ijl}m_{iij}}{m_{il}m_{ij}} \delta_{jk} \delta_{il} + \frac{m_{ijl}m_{ijj}}{m_{jl}m_{ij}} \delta_{ik} \delta_{jl} - \frac{m_{ijl}^2}{m_{il}m_{jl}} \delta_{jk} - \frac{m_{ijl}^2}{m_{il}m_{jl}} \delta_{ik} + \right. \\
&\quad + \frac{m_{ijk}m_{ijj}}{m_{jk}m_{ij}} \delta_{il} \delta_{jk} + \frac{m_{ijk}m_{iij}}{m_{ik}m_{ij}} \delta_{jl} \delta_{ik} - \frac{m_{ijk}^2}{m_{ik}m_{kj}} \delta_{il} - \frac{m_{ijk}^2}{m_{ik}m_{jk}} \delta_{jl} + \\
&\quad + \left( \frac{m_{ijl}}{m_{il}} \right)^2 \delta_{jk} \delta_{ji} \delta_{ik} + \left( \frac{m_{ijk}}{m_{jk}} \right)^2 \delta_{il} \delta_{ij} \delta_{jl} + \\
&\quad \left. + \left( \frac{m_{ijk}}{m_{ik}} \right)^2 \delta_{jl} \delta_{ji} \delta_{il} + \left( \frac{m_{ijl}}{m_{jl}} \right)^2 \delta_{ik} \delta_{ji} \delta_{jk} \right] = \\
&= -\frac{1}{8} \sum_{1 \leq k < l \leq n} \left[ \frac{m_{llk}m_{llk}}{m_{ll}m_{kl}m_{kl}} + \frac{m_{lkk}m_{lkk}}{m_{lk}m_{kk}m_{kl}} + \frac{1}{m_{kk}} \left( \frac{m_{kkl}}{m_{kl}} \right)^2 + \frac{1}{m_{ll}} \left( \frac{m_{llk}}{m_{kl}} \right)^2 - \right. \\
&\quad - \sum_{i=1}^n \left( \frac{m_{ikl}^2}{m_{il}m_{kl}m_{ik}} \right) - \sum_{j=1}^n \left( \frac{m_{jlk}^2}{m_{lk}m_{jk}m_{lj}} \right) - \left( \frac{m_{kkl}m_{kkk}}{m_{kk}m_{kl}m_{kk}} \delta_{kl} + \right. \\
&\quad \left. \left. + \frac{m_{llk}m_{lll}}{m_{ll}m_{lk}m_{ll}} \delta_{lk} \right) \right] = \tag{4.36}
\end{aligned}$$

$$= -\frac{1}{8} \sum_{k,l=1}^n \left( \frac{m_{kkl}^2}{m_{kk}m_{kl}^2} + \frac{m_{kll}^2}{m_{ll}m_{kl}^2} \right) + \frac{1}{4} \sum_{k=1}^n \frac{m_{kkk}^2}{m_{kk}^3} + \frac{1}{4} \sum_{1 \leq k < l \leq n} \sum_{i=1}^n \frac{m_{ikl}^2}{m_{il}m_{kl}m_{ik}} = \quad (4.37)$$

$$= -\frac{1}{8} \sum_{k,l=1}^n \left( \frac{2m_{kkl}^2}{m_{kk}m_{kl}^2} \right) + \frac{1}{4} \sum_{k=1}^n \frac{1}{4} m_{kk} + \frac{1}{4} \sum_{1 \leq i < k < l \leq n} \frac{m_{ikl}^2}{m_{il}m_{kl}m_{ik}} + \frac{1}{4} \sum_{1 \leq k < l \leq n} \frac{m_{llk}^2}{m_{ll}m_{kl}^2} +$$

$$+ \frac{1}{4} \sum_{1 \leq k < l < i \leq n} \frac{m_{ikl}^2}{m_{il}m_{kl}m_{ik}} + \frac{1}{4} \sum_{1 \leq k < l \leq n} \frac{m_{kkl}^2}{m_{kk}m_{kl}^2} + \frac{1}{4} \sum_{1 \leq k < i < l \leq n} \frac{m_{ikl}^2}{m_{il}m_{kl}m_{ik}} = \quad (4.38)$$

$$= \frac{3}{4} \sum_{1 \leq u < v < w \leq n} \frac{m_{uvw}^2}{m_{uv}m_{vw}m_{wu}} + \frac{1}{16} \sum_{k=1}^n m_{kk} - \frac{1}{4} \sum_{k=1}^n \frac{m_{kkk}^2}{m_{kk}^3} = \quad (4.39)$$

$$= \frac{3}{4} \sum_{1 \leq u < v < w \leq n} \frac{m_{uvw}^2}{m_{uv}m_{vw}m_{wu}}. \quad (4.40)$$

Ezzel kiszámoltuk a (4.30) képlet második tagját.

A (4.26) képlet harmadik összege:

$$\sum_{t \in I_2, s \in I_1} \left\langle \tilde{R}(D)(B_s, A_t) \left( \tilde{G}^{(-1)}(D)(A_t) \right), B_s \right\rangle = \quad (4.41)$$

$$= \sum_{1 \leq k < l \leq n} \sum_{i=1}^n \left\langle \tilde{R}(D) \left( \frac{H_{kl}}{\sqrt{2}}, \frac{F_{ii}}{2} \right) \frac{1}{m_{ii}} \frac{F_{ii}}{2}, \frac{H_{kl}}{\sqrt{2}} \right\rangle = \quad (4.42)$$

$$= \sum_{1 \leq k < l \leq n} \sum_{i=1}^n \frac{1}{8m_{ii}} \langle R(D)(H_{kl}, F_{ii}) F_{ii}, H_{kl} \rangle = \quad (4.43)$$

$$\begin{aligned}
&= \sum_{1 \leq k < l \leq n} \sum_{i=1}^n \frac{-1}{32m_{ii}} \left\langle \tilde{G}^{(-1)}(D) \left( d\tilde{G}(D)(H_{kl}) \left( \tilde{G}^{(-1)}(D) \left( d\tilde{G}(D)(F_{ii})(F_{ii}) \right) \right) \right), H_{kl} \right\rangle \\
&+ \sum_{1 \leq k < l \leq n} \sum_{i=1}^n \frac{1}{32m_{ii}} \left\langle \tilde{G}^{(-1)}(D) \left( d\tilde{G}(D)(F_{ii}) \left( \tilde{G}^{(-1)}(D) \left( d\tilde{G}(D)(H_{kl})(F_{ii}) \right) \right) \right), H_{kl} \right\rangle.
\end{aligned} \tag{4.44}$$

A (4.44) képlet első tagja:

$$\sum_{1 \leq k < l \leq n} \sum_{i=1}^n \frac{-1}{32m_{ii}} \left\langle \tilde{G}^{(-1)}(D) \left( d\tilde{G}(D)(H_{kl}) \left( \tilde{G}^{(-1)}(D) \left( d\tilde{G}(D)(F_{ii})(F_{ii}) \right) \right) \right), H_{kl} \right\rangle = \tag{4.45}$$

$$= \sum_{1 \leq k < l \leq n} \sum_{i=1}^n \frac{1}{32m_{ii}} \left\langle \tilde{G}^{(-1)}(D) \left( d\tilde{G}(D)(H_{kl}) \left( \tilde{G}^{(-1)}(D) (4F_{ii}m_{iii}) \right) \right), H_{kl} \right\rangle = \tag{4.46}$$

$$= \sum_{1 \leq k < l \leq n} \sum_{i=1}^n \frac{m_{iii}}{8m_{ii}^2} \left\langle \tilde{G}^{(-1)}(D) (d\tilde{G}(D)(H_{kl})(F_{ii})), H_{kl} \right\rangle = \tag{4.47}$$

$$= \sum_{1 \leq k < l \leq n} \sum_{i=1}^n -\frac{1}{8} \left\langle \tilde{G}^{(-1)}(D) (2H_{ki}m_{kli}\delta_{li} + 2H_{il}m_{kli}\delta_{ki}), H_{kl} \right\rangle = \tag{4.48}$$

$$= -\frac{1}{4} \sum_{1 \leq k < l \leq n} \sum_{i=1}^n \left( \frac{m_{kli}}{m_{ki}} \delta_{li} + \frac{m_{kli}}{m_{li}} \delta_{ki} \right) = \tag{4.49}$$

$$= -\frac{1}{4} \sum_{k,l=1}^n \frac{m_{kll}}{m_{kl}} + \frac{1}{4} \sum_{k=1}^n \frac{m_{kkk}}{m_{kk}} = \tag{4.50}$$

$$= -\frac{1}{4} \sum_{k,l=1}^n \frac{m_{kll}}{m_{kl}} + \frac{1}{8} \sum_{k=1}^n m_{kk}. \tag{4.51}$$

A (4.44) képlet második tagja:

$$\sum_{1 \leq k < l \leq n} \sum_{i=1}^n \frac{1}{32m_{ii}} \left\langle \tilde{G}^{(-1)}(D) \left( d\tilde{G}(D)(F_{ii}) \left( \tilde{G}^{(-1)}(D) \left( d\tilde{G}(D)(H_{kl})(F_{ii}) \right) \right) \right), H_{kl} \right\rangle = \tag{4.52}$$

$$= \sum_{1 \leq k < l \leq n} \sum_{i=1}^n -\frac{1}{16m_{ii}} \times \quad (4.53)$$

$$\begin{aligned} & \times \left\langle \tilde{G}^{(-1)}(D) \left( \frac{m_{kli}}{m_{ki}} \delta_{il} d\tilde{G}(D)(F_{ii})(H_{ki}) + \frac{m_{kli}}{m_{li}} \delta_{ik} d\tilde{G}(D)(F_{ii})(H_{il}) \right), H_{kl} \right\rangle = \\ & = \frac{1}{4} \sum_{1 \leq k < l \leq n} \sum_{i=1}^n \frac{1}{m_{ii}} \left( \frac{m_{kli}m_{iik}}{m_{ik}^2} \delta_{il} + \frac{m_{kli}m_{iil}}{m_{il}^2} \delta_{ik} + 2 \frac{m_{kli}m_{iii}}{m_{ik}m_{ii}} \delta_{ik} \delta_{il} + 2 \frac{m_{kli}m_{iii}}{m_{il}m_{ii}} \delta_{ik} \delta_{il} \right) = \end{aligned} \quad (4.54)$$

$$= \frac{1}{4} \sum_{i=1}^n \left[ \sum_{k,l=1}^n \left( \frac{m_{kli}m_{iik}}{m_{ii}m_{ki}^2} \delta_{il} + 2 \frac{m_{kli}m_{iii}}{m_{ik}m_{ii}^2} \delta_{ik} \delta_{il} \right) - \right. \quad (4.55)$$

$$\left. - \sum_{k=1}^n \left( \frac{m_{kki}m_{iik}}{m_{ii}m_{ki}^2} \delta_{ik} + 2 \frac{m_{kki}m_{iii}}{m_{ik}m_{ii}^2} \delta_{ik} \right) \right] =$$

$$= \frac{1}{4} \sum_{i=1}^n \left[ \sum_{k=1}^n \frac{m_{iik}^2}{m_{ii}m_{ki}^2} - \frac{m_{iii}^2}{m_{ii}^3} \right] = \quad (4.56)$$

$$= \frac{1}{4} \sum_{k,l=1}^n \frac{m_{kkl}^2}{m_{kk}m_{kl}^2} - \frac{1}{4} \sum_{k=1}^n \frac{m_{kkk}^2}{m_{kk}^3}. \quad (4.57)$$

Ezzel kiszámoltuk a (4.30) képlet harmadik tagját.

A (4.30) képlet negyedik és ötödik összege:

Szimmetriaokokból a negyedik- illetve ötödik összeg megegyezik a második- illetve harmadik összeggel.

A (4.30) képlet hatodik összege:

$$\sum_{t,s \in I_1} \left\langle \tilde{R}(D)(B_s, B_t) \left( \tilde{G}^{(-1)}(D)(B_t) \right), B_s \right\rangle = \quad (4.58)$$

$$= \sum_{\substack{1 \leq i < j \leq n \\ 1 \leq k < l \leq n}} \left\langle \tilde{R}(D) \left( \frac{H_{kl}}{\sqrt{2}}, \frac{H_{ij}}{\sqrt{2}} \right) \frac{1}{m_{ij}} \frac{H_{ij}}{\sqrt{2}}, \frac{H_{kl}}{\sqrt{2}} \right\rangle = \quad (4.59)$$

$$= \sum_{\substack{1 \leq i < j \leq n \\ 1 \leq k < l \leq n}} \frac{1}{4m_{ij}} \langle \tilde{R}(D)(H_{kl}, H_{ij})H_{ij}, H_{kl} \rangle = \quad (4.60)$$

$$= \sum_{\substack{1 \leq i < j \leq n \\ 1 \leq k < l \leq n}} \frac{-1}{16m_{ij}} \left\langle \tilde{G}^{(-1)}(D) \left( d\tilde{G}(D)(H_{kl}) \left( \tilde{G}^{(-1)}(D) \left( d\tilde{G}(D)(H_{ij})(H_{ij}) \right) \right) \right), H_{kl} \right\rangle + \quad (4.61)$$

$$+ \sum_{\substack{1 \leq i < j \leq n \\ 1 \leq k < l \leq n}} \frac{1}{16m_{ij}} \left\langle \tilde{G}^{(-1)}(D) \left( d\tilde{G}(D)(H_{ij}) \left( \tilde{G}^{(-1)}(D) \left( d\tilde{G}(D)(H_{kl})(H_{ij}) \right) \right) \right), H_{kl} \right\rangle .$$

A (4.61) képlet első tagja:

$$\sum_{\substack{1 \leq i < j \leq n \\ 1 \leq k < l \leq n}} \frac{1}{16m_{ij}} \left\langle \tilde{G}^{(-1)}(D) \left( d\tilde{G}(D)(H_{kl}) \left( \tilde{G}^{(-1)}(D) (F_{ii}m_{ij} + m_{ij}F_{jj}) \right) \right), H_{kl} \right\rangle = \quad (4.62)$$

$$= \sum_{\substack{1 \leq i < j \leq n \\ 1 \leq k < l \leq n}} \frac{1}{16m_{ij}} \left\langle \tilde{G}^{(-1)}(D) \left( d\tilde{G}(D)(H_{kl}) \left( \frac{m_{ij}}{m_{ii}} F_{ii} + \frac{m_{ij}}{m_{jj}} F_{jj} \right) \right), H_{kl} \right\rangle = \quad (4.63)$$

$$= \sum_{\substack{1 \leq i < j \leq n \\ 1 \leq k < l \leq n}} -\frac{1}{16m_{ij}} \frac{m_{ij}}{m_{ii}} \left\langle \tilde{G}^{(-1)}(D) (2\delta_{li}m_{kli}H_{ki} + 2\delta_{ki}m_{ikl}H_{il}), H_{kl} \right\rangle + \quad (4.64)$$

$$-\frac{1}{16m_{ij}} \frac{m_{ij}}{m_{jj}} \left\langle \tilde{G}^{(-1)}(D) (2\delta_{lj}m_{klj}H_{kj} + 2\delta_{kj}m_{klj}H_{jl}), H_{kl} \right\rangle =$$

$$= -\sum_{\substack{1 \leq i < j \leq n \\ 1 \leq k < l \leq n}} \left[ \frac{1}{16m_{ij}} \frac{m_{ij}}{m_{ii}} \left( 2\frac{m_{kli}}{m_{ki}} \delta_{li} \langle F_{ki}, F_{kl} \rangle + 2\frac{m_{kli}}{m_{li}} \delta_{ki} \langle F_{li}, F_{kl} \rangle \right) + \quad (4.65)$$

$$+ \frac{1}{16m_{ij}} \frac{m_{ij}}{m_{jj}} \left( 2\frac{m_{klj}}{m_{kj}} \delta_{lj} \langle F_{kj}, F_{kl} \rangle + 2\frac{m_{klj}}{m_{lj}} \delta_{kj} \langle F_{lj}, F_{kl} \rangle \right) \right] .$$

Ez megegyezik a (2.176) képlettel, vagyis a (4.61) képlet első tagját a (2.180) képlet adja meg.

$$= -\frac{1}{2} \left[ \sum_{1 \leq u < v < w \leq n} \left( \frac{m_{vvw}m_{uvw}}{m_{vu}m_{vv}m_{vw}} + \frac{m_{uuv}m_{uuv}}{m_{uu}m_{uv}m_{uw}} + \frac{m_{wvu}m_{wvw}}{m_{wu}m_{wv}m_{ww}} \right) + \right. \quad (4.66)$$

$$\left. + \sum_{1 \leq k < l \leq n} \left( \frac{m_{kll}^2}{m_{kl}^2 m_{ll}} + \frac{m_{kkl}^2}{m_{kl}^2 m_{kk}} \right) \right] \quad (4.67)$$

A (4.61) képlet második tagja:

$$\sum_{\substack{1 \leq i < j \leq n \\ 1 \leq k < l \leq n}} \frac{1}{16m_{ij}} \left\langle \tilde{G}^{(-1)}(D) \left( d\tilde{G}(D)(H_{ij}) \left( \tilde{G}^{(-1)}(D) \left( d\tilde{G}(D)(H_{kl})(H_{ij}) \right) \right) \right), H_{kl} \right\rangle = \quad (4.68)$$

$$= \sum_{\substack{1 \leq i < j \leq n \\ 1 \leq k < l \leq n}} -\frac{1}{16m_{ij}} \left\langle \tilde{G}^{(-1)}(D) \left( d\tilde{G}(D)(H_{ij}) \left( \tilde{G}^{(-1)}(D) (F_{il}m_{ijl}\delta_{jk} + F_{jk}m_{ijk}\delta_{il} + \right. \right. \right. \quad (4.69)$$

$$\left. \left. \left. - F_{ik}m_{ijk}\delta_{jl} - F_{jl}m_{ijl}\delta_{ik}) \right) \right), H_{kl} \right\rangle =$$

$$= \sum_{\substack{1 \leq i < j \leq n \\ 1 \leq k < l \leq n}} -\frac{1}{16m_{ij}} \left\langle \tilde{G}^{(-1)}(D) \left( d\tilde{G}(D)(H_{ij}) (F_{il}) \frac{m_{ijl}}{m_{il}} \delta_{jk} + d\tilde{G}(D)(H_{ij}) (F_{jk}) \frac{m_{ijk}}{m_{jk}} \delta_{il} - \right. \right. \quad (4.70)$$

$$\left. \left. - d\tilde{G}(D)(H_{ij}) (F_{ik}) \frac{m_{ijk}}{m_{ik}} \delta_{jl} - d\tilde{G}(D)(H_{ij}) (F_{jl}) \frac{m_{ijl}}{m_{jl}} \delta_{ik} \right), H_{kl} \right\rangle =$$

$$\begin{aligned}
&= \sum_{\substack{1 \leq i < j \leq n \\ 1 \leq k < l \leq n}} -\frac{1}{16m_{ij}} \times \\
&\quad \times \left\langle \tilde{G}^{(-1)}(D) \left( \frac{m_{ijl}}{m_{il}} \delta_{jk} (H_{ii} m_{ijj} \delta_{jl} - H_{il} m_{ijl} \delta_{ji} - H_{ij} m_{iji} \delta_{il} + H_{jl} m_{ijl} \delta_{ii}) + \right. \right. \\
&\quad + \frac{m_{ijk}}{m_{jk}} \delta_{il} (H_{ji} m_{ijj} \delta_{jk} - H_{ik} m_{ijk} \delta_{jj} - H_{jj} m_{ijj} \delta_{ik} + H_{jk} m_{ijk} \delta_{ij}) + \\
&\quad + \frac{m_{ijk}}{m_{ik}} \delta_{jl} (H_{ii} m_{ijj} \delta_{jk} - H_{ik} m_{ijk} \delta_{ji} + H_{ji} m_{iji} \delta_{ik} + H_{jk} m_{ijk} \delta_{ii}) + \\
&\quad \left. \left. + \frac{m_{ijl}}{m_{jl}} \delta_{ik} (H_{ij} m_{ijj} \delta_{jl} - H_{il} m_{ijl} \delta_{jj} - H_{jj} m_{ijj} \delta_{il} + H_{jl} m_{ijl} \delta_{ij}) \right), F_{kl} \right\rangle = \\
&\hspace{15em} (4.71)
\end{aligned}$$

$$\begin{aligned}
&= -\frac{1}{8} \sum_{\substack{1 \leq i < j \leq n \\ 1 \leq k < l \leq n}} \frac{1}{m_{ij}} \left[ \frac{m_{ijl} m_{ijj}}{m_{il} m_{ij}} \delta_{jk} \delta_{il} + \frac{m_{ijl} m_{ijj}}{m_{jl} m_{ij}} \delta_{ik} \delta_{jl} - \frac{m_{ijl}^2}{m_{il} m_{jl}} \delta_{jk} - \frac{m_{ijl}^2}{m_{il} m_{jl}} \delta_{ik} + \right. \\
&\quad + \frac{m_{ijk} m_{ijj}}{m_{jk} m_{ij}} \delta_{il} \delta_{jk} + \frac{m_{ijk} m_{ijj}}{m_{ik} m_{ij}} \delta_{jl} \delta_{ik} - \frac{m_{ijk}^2}{m_{ik} m_{kj}} \delta_{il} - \frac{m_{ijk}^2}{m_{ik} m_{jk}} \delta_{jl} + \\
&\quad + \left( \frac{m_{ijl}}{m_{il}} \right)^2 \delta_{jk} \delta_{ji} \delta_{ik} + \left( \frac{m_{ijk}}{m_{jk}} \right)^2 \delta_{il} \delta_{ij} \delta_{jl} + \\
&\quad \left. + \left( \frac{m_{ijk}}{m_{ik}} \right)^2 \delta_{jl} \delta_{ji} \delta_{il} + \left( \frac{m_{ijl}}{m_{jl}} \right)^2 \delta_{ik} \delta_{ji} \delta_{jk} \right]. \\
&\hspace{15em} (4.72)
\end{aligned}$$

Ez megegyezik a (4.36) képlettel, ezért az alábbi végeredményt kapjuk.

$$\frac{1}{4} \sum_{1 \leq u < v < w \leq n} \frac{3m_{uvw}^2}{m_{uv} m_{vw} m_{wu}} + \frac{1}{4} \sum_{1 \leq k < l \leq n} \left( \frac{m_{kkl}^2}{m_{kk} m_{kl}^2} + \frac{m_{kll}^2}{m_{ll} m_{kl}^2} \right) \quad (4.73)$$

Ezzel kiszámoltuk a (4.30) képlet hatodik tagját.

Az  $(\tilde{\mathcal{M}}_n^+, \tilde{K}^{(n),f})$  skalárgörbület kiszámításához az alábbi tagokat kell figyelembe venni a (4.26) képletnek megfelelően:

1-szer OFF-OFF ( $\mathbb{R}$ - $\mathbb{R}$ ): a (2.180) és a (2.188) képlet összege;

2-szer OFF-OFF ( $\mathbb{R}$ - $\mathbb{C}$  és  $\mathbb{C}$ - $\mathbb{R}$ ) a (4.32) és a (4.40) képlet összege;

1-szer OFF-OFF ( $\mathbb{C}$ - $\mathbb{C}$ ): a (4.66) és a (4.73) képlet összege;

2-szer OFF-DIAG, DIAG-OFF ( $\mathbb{R}$ - $\mathbb{R}$ ): a (2.198) és a (2.204) képlet összege;

2-szer OFF-DIAG, DIAG-OFF ( $\mathbb{R}$ - $\mathbb{C}$ ): a (4.51) és a (4.57) képlet összege;

1-szer DIAG-DIAG ( $\mathbb{R}$ - $\mathbb{C}$ ): a (2.226) képlet.

Ezek alapján a skalárgörbületre

$$\begin{aligned} \tilde{\text{Scal}}_D = & \sum_{1 \leq u < v < w \leq n} \left( \frac{3m_{uvw}^2}{m_{uv}m_{vw}m_{wu}} - \frac{2m_{uuv}m_{uuw}}{m_{uu}m_{uv}m_{uw}} - \frac{2m_{uuv}m_{vvw}}{m_{vu}m_{vv}m_{vw}} - \frac{2m_{uww}m_{vww}}{m_{wu}m_{wv}m_{ww}} \right) + \\ & + \frac{1}{2} \sum_{1 \leq k < l \leq n} \frac{1}{m_{kl}^2} \left( \frac{m_{kkl}^2}{m_{kk}} + \frac{m_{kll}^2}{m_{ll}} \right) - \sum_{1 \leq k < l \leq n} \frac{m_{kkl} + m_{kll}}{m_{kl}} \end{aligned} \quad (4.74)$$

teljesül. Az

$$\sum_{1 \leq u < v < w \leq n} \frac{3m_{uvw}^2}{m_{uv}m_{vw}m_{wu}} = \sum_{1 \leq k < l \leq n} \sum_{j=1}^n \frac{m_{jkl}^2}{m_{jl}m_{lk}m_{kj}} - \sum_{1 \leq k < l \leq n} \frac{1}{m_{kl}^2} \left( \frac{m_{kkl}^2}{m_{kk}} + \frac{m_{kll}^2}{m_{ll}} \right) \quad (4.75)$$

$$\begin{aligned} & \sum_{1 \leq u < v < w \leq n} \left( \frac{2m_{uuv}m_{uuw}}{m_{uu}m_{uv}m_{uw}} + \frac{2m_{uuv}m_{vvw}}{m_{vu}m_{vv}m_{vw}} + \frac{2m_{uww}m_{vww}}{m_{wu}m_{wv}m_{ww}} \right) = \\ & = \sum_{1 \leq k < l \leq n} \sum_{j=1}^n \left( \frac{m_{kkl}m_{kkj}}{m_{kj}m_{kk}m_{kl}} + \frac{m_{llj}m_{kll}}{m_{lj}m_{ll}m_{jl}} \right) - \sum_{1 \leq k < l \leq n} \frac{1}{m_{kl}^2} \left( \frac{m_{kkl}^2}{m_{kk}} + \frac{m_{kll}^2}{m_{ll}} \right) - \\ & - \frac{1}{2} \sum_{1 \leq k < l \leq n} \frac{m_{kkl} + m_{kll}}{m_{kl}} \end{aligned} \quad (4.76)$$

azonosságok felhasználásával alakítjuk tovább a skalárgörbületre kapott (4.74) kifejezést.



$$\begin{aligned} \tilde{\text{Scal}}(D) = & \sum_{1 \leq k < l \leq n} \frac{1}{m_{kl}} \sum_{j=1}^n \left( \frac{m_{jkl}^2}{m_{jl}m_{jk}} - \frac{m_{kk}m_{kkj}}{m_{kk}m_{kj}} - \frac{m_{lj}m_{kll}}{m_{lj}m_{ll}} \right) + \\ & + \frac{1}{2} \sum_{1 \leq k < l \leq n} \frac{1}{m_{kl}^2} \left( \frac{m_{kk}^2}{m_{kk}} + \frac{m_{ll}^2}{m_{ll}} \right) - \frac{1}{2} \sum_{1 \leq k < l \leq n} \frac{m_{kk} + m_{ll}}{m_{kl}} \end{aligned} \quad (4.77)$$

$$\tilde{\text{Scal}}(D) = \sum_{j,k,l=1}^n \left( \frac{1}{2} \frac{m_{jkl}^2}{m_{jl}m_{jk}} - \frac{m_{kk}m_{kkj}}{m_{kk}m_{kj}} \right) + \frac{1}{8} \sum_{k=1}^n m_{kk} \quad (4.78)$$

Ezek alapján

$$\tilde{\text{Scal}}_D = \sum_{\substack{j,k,l=1 \\ |\{j,k,l\}| > 1}}^n \left( \frac{1}{2} \frac{m_{jkl}^2}{m_{jk}m_{kl}m_{lj}} - \frac{m_{kk}m_{kkj}}{m_{kk}m_{kl}m_{kj}} \right) \quad (4.79)$$

teljesül. Ezzel meghatároztuk komplex állapotter esetén az  $(\tilde{\mathcal{M}}_n^+, \tilde{K}^{(n),f})$  sokaság skalárgörbületét a  $D = \sum_{i=1}^n \lambda_i E_{ii}$  pontban.

Az eddigi számolásokat foglalja össze a következő tétel.

**4.2. Tétel.** *Legyen  $f \in \mathcal{F}_{[0,\infty]}^{(S,n)}$  tetszőleges függvény, és jelölje  $K^{(n),f}$  az  $f$  függvény által indukált monoton metrikát az  $\mathcal{M}_n^+$  sokaságon. Legyen  $D \in \mathcal{M}_n^+$  tetszőleges állapot, melynek  $\lambda_1, \dots, \lambda_n$  a sajátértékei. A  $(\mathcal{M}_n^+, K^{(n),f})$  Riemann-sokaság skalárgörbülete a  $D$  pontban, valós állapotter esetén*

$$\begin{aligned} \text{Scal}_{\mathbb{R}}(D) = & \frac{1}{4} \sum_{\substack{j,k,l=1 \\ |\{j,k,l\}| > 1}}^n \left( \frac{1}{2} \frac{m_{jkl}^2}{m_{jk}m_{kl}m_{lj}} - \frac{m_{kk}m_{kkj}}{m_{kk}m_{kl}m_{kj}} \right) + \frac{1}{4} \sum_{\substack{k,l=1 \\ k \neq l}}^n \left( \frac{1}{2} \frac{m_{kkl}^2}{m_{kl}^2 m_{kk}} - \frac{m_{kkl}m_{kll}}{m_{kl}^2 m_{ll}} \right) + \\ & + \frac{(n-1)(n+2)(n^2+n-4)}{16}, \end{aligned} \quad (4.80)$$

komplex állapotter esetén pedig

$$\text{Scal}(D) = \sum_{\substack{j,k,l=1 \\ |\{j,k,l\}| > 1}}^n \left( \frac{1}{2} \frac{m_{jkl}^2}{m_{jk}m_{kl}m_{lj}} - \frac{m_{kk}m_{kkj}}{m_{kk}m_{kl}m_{kj}} \right) + \frac{(n^2-1)(n^2-2)}{16}. \quad (4.81)$$

## 4.2. A skalárgörbület monotonitásáról

Az elemi térfogati forma sorfejtését felhasználva Petz megmutatta, hogy az állapotter skálárgörbülete az állapot statisztikai bizonytalanságával, megkülönböztetetheletlenségével van szoros kapcsolatban [84]. Kvantummechanikai tapasztalatok alapján várhatjuk, hogy a kevertebb állapotok kevésbé megkülönböztethetők. A matematika nyelvén ez azt jelenti, hogy a fizikailag, statisztikailag releváns Riemann-metrikákból származó skálárgörbületnek egyfajta monotonitási tulajdonsággal kell rendelkeznie: ha a  $D_1$  állapot kevertebb mint a  $D_2$  állapot, akkor a  $\text{Scal}(D_2) < \text{Scal}(D_1)$  egyenlőtlenségnek kell teljesülnie.

Fizikai szempontból a Kubo–Mori-metrika kitüntetett szerepet játszik az állapotterén. Petz Dénes sejtése [81] erre a metrikára és az előbbi észrevételre vonatkozik.

**Petz sejtése:** *A Kubo–Mori-metrikából származó skálárgörbület monoton csökkenő az állapotok majorizációjára nézve, vagyis*

$$\text{ha } D_1 \prec D_2, \text{ akkor } \text{Scal}(D_1) \geq \text{Scal}(D_2) \quad (4.82)$$

*teljesül.*

Különböző monoton Riemann-metrikák esetén is vizsgáljuk a későbbiekben a sejtésben megfogalmazott monotonitási tulajdonságot, ezért külön elnevezést vezetünk be rá.

**4.1. Definíció.** Azt mondjuk, hogy az  $f : \mathcal{M}_n^+ \rightarrow \mathbb{R}$  függvény a *majorizációra nézve monoton csökkenő*, ha minden  $D_1, D_2 \in \mathcal{M}_n^+$  állapotra  $D_1 \prec D_2$  esetén  $f(D_1) \geq f(D_2)$  teljesül.

Petz a  $2 \times 2$ -es mátrixok állapotterén be is bizonyította ezt a sejtését [81]. Ezen túlmenően eddig csak numerikus szimulációkat végeztek a sejtéssel kapcsolatban, amelyek mind megerősítették azt. Érdeemes megjegyezni, hogy a klasszikus esetben a sejtés teljesül, hiszen a (2.50) formula alapján a skálárgörbület ott állandó. A továbbiakban a sejtés bizonyításában Andai eddig elért eredményeit mutatjuk be [4] alapján.

### 4.2.1. Petz sejtésének átfogalmazása

Legyenek  $A$  és  $B$  olyan állapotok, melyekre  $A \prec B$  teljesül. Ekkor a 3.5. tétel alapján létezik olyan invertálható állapotokból álló  $(C_z)_{z=1,\dots,d}$  véges sorozat, amellyel

$$A = C_1 \prec C_2 \prec \dots \prec C_d = B \quad (4.83)$$

teljesül, és  $C_z$  sajátértékeiből  $T$ -transzformációval megkaphatjuk  $C_{z-1}$  sajátértékeit, minden  $z = 2, \dots, d$  esetén. A (4.82) sejtés igazolásához elég megmutatni, hogy minden  $z = 2, \dots, d$  esetén

$$\text{Scal}(C_{z-1}) \geq \text{Scal}(C_z) \quad (4.84)$$

fennáll. A skalárgörbület csak a sajátértékektől függ, ezért a fenti egyenlőtlenséghez elég igazolni, hogy ha  $(x_1, \dots, x_n)$  jelöli a  $C_z$  állapot sajátértékeit, akkor minden  $1 \leq k < l \leq n$  esetén a

$$\left[0, \frac{1}{2}\right] \rightarrow \mathbb{R} \quad t \mapsto \text{Scal}(x_1, \dots, x_{k-1}, tx_k + (1-t)x_l, \dots, x_{l-1}, (1-t)x_k + tx_l, \dots, x_n) \quad (4.85)$$

függvény monoton növvő. Petz sejtésének ezen átfogalmazását fogjuk részletesen vizsgálni.

#### 4.2.2. Skalárgörbület a Kubo-Mori metrikánál

Az  $\mathcal{M}_n^+$  téren a  $K_{\text{KM}}^{(n)}$  Kubo-Mori-metrikát a  $D \in \mathcal{M}_n^+$  állapotban és  $X, Y \in T_D \mathcal{M}_n^+$  érintővektorokon a

$$G_D(X, Y) = \int_0^\infty \text{Tr}((D+t)^{-1}X(D+t)^{-1}Y) dt \quad (4.86)$$

képlet határozza meg. Ezzel a metrikával számolva az alábbi kifejezések adódnak a (2.127) és (2.145) képlettel értelmezett  $m$ -függvényekre.

$$m(x, y) = \int_0^\infty \frac{1}{(x+t)(y+t)} dt \quad m(x, y, z) = \int_0^\infty \frac{1}{(x+t)(y+t)(z+t)} dt \quad (4.87)$$

A  $D = \sum_{k=1}^{n=1} \lambda_k E_{kk}$  alakú  $D \in \mathcal{M}_n^+$  állapotban az  $(F_{ij})_{1 \leq i < j \leq n}, (H_{ij})_{1 \leq i < j \leq n}$  bázisvektorok skalárszorzata az érintőtérben a 3.15. tétel szerint az alábbi.

$$\begin{aligned} \text{Ha } 1 \leq i < j \leq n, 1 \leq k < l \leq n, \text{ akkor:} & \begin{cases} G_D(H_{ij}, H_{kl}) = \delta_{ik} \delta_{jl} 2m(\lambda_i, \lambda_j) \\ G_D(F_{ij}, F_{kl}) = \delta_{ik} \delta_{jl} 2m(\lambda_i, \lambda_j) \\ G_D(H_{ij}, F_{kl}) = 0, \end{cases} \\ \text{ha } 1 \leq i < j \leq n, 1 \leq k \leq n, \text{ akkor:} & G_D(H_{ij}, F_{kk}) = G(F_{ij}, F_{kk}) = 0, \\ \text{ha } 1 \leq i \leq n, 1 \leq k \leq n, \text{ akkor:} & G_D(F_{ii}, F_{kk}) = \delta_{ik} 4m(\lambda_i, \lambda_i). \end{aligned} \quad (4.88)$$

Legyenek  $x, y, z$  és  $\mu$  pozitív számok,  $x \neq y$ , ekkor

$$m(x, y) = \frac{\log x - \log y}{x - y}, \quad m(x, y, z) = \frac{m(x, z) - m(y, z)}{x - y}, \quad (4.89)$$

$$m(x, x) = \frac{1}{x} \quad m(x, x, x) = \frac{1}{2x^2}, \quad m(x, y) = \frac{1}{\mu} m(\mu x, \mu y),$$

$$m(x, x, y) = \frac{m(x, y) - \frac{1}{x}}{x - y}, \quad m(x, y, z) = \frac{1}{\mu^2} m(\mu x, \mu y, \mu z)$$

teljesül.

Az  $m$  függvények teljesítik az

$$\frac{1}{m(x, y)} \left( \frac{m(x, x, y)}{m(x, x)} + \frac{m(x, y, y)}{m(y, y)} \right) = 1 \quad (4.90)$$

azonosságot, melyet gyakran említés nélkül használunk fel a számolás során.

A rövidebb írásmód kedvéért vezessük be az alábbi függvényeket

$$\varphi(x, y, z) = \frac{1}{2} \frac{m(x, y, z)^2}{m(x, y)m(y, z)m(z, x)} - \frac{m(y, y, x)m(y, y, z)}{m(y, x)m(y, y)m(y, z)}, \quad (4.91)$$

$$v(x, y) = \frac{1}{2} \frac{m(x, x, y)^2}{m(x, y)^2 m(x, x)} - \frac{m(x, x, y)m(y, y, x)}{m(x, y)^2 m(y, y)}.$$

A definícióból és a (4.89) azonosságból a

$$\varphi(\mu x, \mu y, \mu z) = \frac{1}{\mu} \varphi(x, y, z) \quad v(\mu x, \mu y) = \frac{1}{\mu} v(x, y). \quad (4.92)$$

skalátulajdonság adódik.

A komplex mátrixokból álló állapot térnek a skalárgömbülete a  $D$  állapotban, ha  $D$  sajátértékei  $\lambda_1, \dots, \lambda_n$ , a (4.80) egyenlet alapján

$$\text{Scal}(D) = \sum_{\substack{j, k, l=1 \\ |\{j, k, l\}| > 1}}^n \varphi(\lambda_j, \lambda_k, \lambda_l) + \frac{(n^2 - 1)(n^2 - 2)}{16}, \quad (4.93)$$

ahol  $|\{j, k, l\}| > 1$  azt jelenti, hogy az  $i, j, k$  indexek nem mind azonosak. Valós mátrixok esetén a skalárgömbület a  $D$  állapotban a (4.81) alapján

$$\text{Scal}_{\mathbb{R}} D = \frac{1}{4} \sum_{\substack{j, k, l=1 \\ |\{j, k, l\}| > 1}}^n \varphi(\lambda_j, \lambda_k, \lambda_l) + \frac{1}{4} \sum_{k, l=1}^n v(\lambda_k, \lambda_l) + \frac{(n-1)(n+2)(n^2+n-4)}{16}. \quad (4.94)$$

Mivel a jelen részben a skalárgörbület függvény változását vizsgáljuk, ezért a skalárgörbület kifejezés utolsó, csak  $n$ -től függő tagját elhagyjuk.

### 4.2.3. A Petz sejtésében elért eredmények

Legyenek  $A$  és  $B$  olyan állapotok, melyekre  $A \prec B$  teljesül, és a csökkenő sorrendbe rendezett sajátértékeik két kivétellel megegyeznek. Ekkor az állapotok sajátértékeit a  $(\lambda_1, \dots, a, \dots, b, \dots, \lambda_n)$  és  $(\lambda_1, \dots, a-x, \dots, b+x, \dots, \lambda_n)$  alakban írhatjuk fel, ahol  $x \leq \frac{a-b}{2}$ .

A skalárgörbületet adó (4.93) képletben szereplő tagokat csoportosítsuk az alábbiak szerint.

1. Azok a tagok, amelyekben csak  $a$  és  $b$  szerepel

$$\alpha(a, b) := 2\varphi(a, a, b) + 2\varphi(b, b, a) + \varphi(a, b, a) + \varphi(b, a, b) . \quad (4.95)$$

2. Azok a tagok, amelyekben  $a, b$  és egy másik sajátérték is szerepel

$$\beta_{1,k}(a, b) := 2\varphi(a, a, \lambda_k) + 2\varphi(b, b, \lambda_k) + \varphi(a, \lambda_k, a) + \varphi(b, \lambda_k, b) , \quad (4.96)$$

$$\beta_{2,k}(a, b) := 2\varphi(a, b, \lambda_k) + 2\varphi(b, a, \lambda_k) + \varphi(a, \lambda_k, b) + \varphi(b, \lambda_k, a) . \quad (4.97)$$

3. Azok a tagok, amelyekben vagy  $a$  vagy  $b$  pontosan egyszer szerepel

$$\begin{aligned} \gamma_{kl}(a, b) := & \varphi(a, \lambda_k, \lambda_l) + \varphi(b, \lambda_k, \lambda_l) + \varphi(\lambda_k, a, \lambda_l) + \\ & + \varphi(\lambda_k, b, \lambda_l) + \varphi(\lambda_k, \lambda_l, a) + \varphi(\lambda_k, \lambda_l, b) . \end{aligned} \quad (4.98)$$

4. Az  $a$  és  $b$  nélküli tagok

$$\delta_{jkl} := \varphi(\lambda_j, \lambda_k, \lambda_l) . \quad (4.99)$$

Ezután a skalárgörbületet a

$$\text{Scal}(D) = \alpha(a, b) + \sum_{k=1}^n (\beta_{1,k}(a, b) + \beta_{2,k}(a, b)) + \sum_{k,l=1}^n \gamma_{kl}(a, b) + \sum_{\substack{j,k,l=1 \\ |\{j,k,l\}|>1}}^n \delta_{jkl} \quad (4.100)$$

egyenlet adja meg.

A sejtés bizonyításának az egyik lehetséges módja, hogy megmutatjuk, hogy a fenti képletben szereplő összeadandó tagok külön-külön monotonak a majorizációra nézve. Ennek érdekében a (4.85) képlet alapján vezetjük be az alábbi fogalmat.

**4.2. Definíció.** Az  $f : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$  függvényről azt mondjuk, hogy *m-monoton*, ha minden olyan  $a$  és  $b$  paraméter esetén, amelyre  $1 > a > b > 0$ , az  $f(a-x, b+x)$  függvény szigorúan monoton növekvő függvénye  $x$ -nek az  $x \in [0, \frac{a+b}{2}]$  intervallumon.

Dittmann a  $\varphi$  függvény szimmetrizált alakját használta, és abból származtatta a tagok fent említett csoportosítását [25]. Numerikus szimulációkat végzett a szimmetrizált  $\varphi$  függvény monotonitásának a vizsgálatára, és meggyőző ábrákon foglalta össze a szimulációk végeredményét. Azonban korrekt matematikai indokkal nem tudta igazolni a szimulációk eredményét.

A sejtés bizonyításában elért részeredmények alapja az entrópiafüggvény deriváltjainak a viselkedésének a megértése. Ugyanis az entrópia második deriváltja szolgáltatja a metrikát, melynek a második deriváltjától (is) függ a skalárgörbület, továbbá a sejtésben elért részeredmények bizonyítása a skalárgörbület deriváltjainak a viselkedésén alapul. (Például az egyik esetben –közvetett módon– a skalárgörbület 14. deriváltjának a viselkedésén múlik a bizonyítandó egyenlőtlenség, vagyis az entrópia 18. deriváltja is előkerül!) A deriváltak tulajdonságainak még pontosabb ismerete valószínűleg elegendő lenne a sejtés bizonyításához.

**4.3. Tétel.** *A skalárgörbületben szereplő  $\alpha$  függvény m-monoton.*

*Bizonyítás.* A (4.89) és (4.91) azonosság alapján

$$\varphi(a, b, a) + \varphi(b, a, b) = -\frac{1}{2(a+b)} \frac{1+b/a}{m(1, b/a)^2} (m(1, 1, b/a)^2 + (b/a)m(b/a, b/a, 1)^2) \quad (4.101)$$

és

$$2\varphi(a, a, b) + 2\varphi(b, b, a) = -\frac{1}{a+b} (1+b/a)^2 \frac{m(1, 1, b/a)m(b/a, b/a, 1)}{n(1, b/a)^2}. \quad (4.102)$$

Az  $m$  függvény speciális értékei (4.89) és a fenti egyenletek alapján

$$\alpha(a-x, b+x) = \frac{-1}{a+b} \cdot \left( -\frac{1}{2} \frac{(1+c(x))^2}{(1-c(x))^2} + \frac{(1+c(x))(1+c(x)^2)}{c(x)(c(x)-1)\log c(x)} - \frac{1}{2} \frac{(1+c(x))^2}{c(x)\log^2 c(x)} \right), \quad (4.103)$$

ahol

$$c(x) = \frac{b+x}{a-x}. \quad (4.104)$$

Mivel a  $c$  függvény monoton növekvő, a tétel bizonyításához elég belátni, hogy a

$$\tau(c) = -\frac{1}{2} \frac{(1+c)^2}{(1-c)^2} + \frac{(1+c)(1+c^2)}{c(c-1)\log c} - \frac{1}{2} \frac{(1+c)^2}{c\log^2 c} \quad (4.105)$$

függvény csökkenő az  $I = [0, 1]$  intervallumon. Annak igazolásához, hogy a

$$\tau'(c) = \frac{(c+1)^2}{c^2 \log^3 c} - \frac{(c+1)(3c^2 - 2c + 3)}{c^2(c-1) \log^2 c} + \frac{c^4 - 2c^3 - 2c^2 - 2c + 1}{c^2(c-1)^2 \log c} + \frac{2c+2}{(c-1)^3} \quad (4.106)$$

függvény csak negatív értékeket vesz fel, vagy vele ekvivalens módon, hogy a

$$\tau_1(c) = (c-1)^3 \log^3 c \cdot \tau'(c) \quad (4.107)$$

függvény csak negatív értékeket vesz fel az  $I$  intervallumon, elég belátni, hogy

$$(a) \tau_1'(c) > 0 \text{ az } I\text{-n és } \lim_{c \rightarrow 1} \tau_1(c) = 0.$$

A határértéket könnyen lehet igazolni, és az (a) állítás első része pedig következik az alábbi (b) állításból:

$$(b) \tau_2(c) > 0 \text{ az } I\text{-n és } \lim_{c \rightarrow 1} \tau_1'(c) = 0,$$

ahol

$$\tau_2(c) = \frac{c^4}{(c-1)(c^2 - c + 1)(c^2 + c + 1)} \cdot \tau_1'(c). \quad (4.108)$$

A  $\tau_1'$  függvényt behelyettesítve az előző képletbe

$$\tau_2(c) = 6 \log^2 c + \frac{(c-1)(c+1)(c^2 - 12c + 1)}{(c^2 + c + 1)(c^2 - c + 1)} \cdot \log c + \frac{(c+1)^2(c-1)^2}{2(c^2 + c + 1)^2(c^2 - c + 1)^2} \quad (4.109)$$

adódik. A (b) állításban szereplő határérték megint könnyen ellenőrizhető, és a  $\tau_2$  függvény pozitivitása pedig következik az alábbi (c) állításból:

$$(c) \tau_3(c) < 0 \text{ az } I \text{ intervallumon és } \lim_{c \rightarrow 1} \tau_2(c) = 0,$$

ahol

$$\tau_3(c) = c(c^2 + c + 1)^2(c^2 - c + 1)^2 \cdot \tau_2'(c). \quad (4.110)$$

Egyszerűsítések után adódik  $\tau_3$  explicit alakja

$$\begin{aligned} \tau_3(c) = & 2(6c^8 + 6c^7 + 13c^6 - 24c^5 + 22c^4 - 24c^3 + 13c^2 + 6c + 6) \log c \\ & + (c^8 - 12c^7 + 4c^6 - 4c^2 + 12c - 1). \end{aligned} \quad (4.111)$$

A (c) állításban szereplő határérték könnyen ellenőrizhető és a  $\tau_3$  függvény negativitása következik az alábbi (d) állításból

$$(d) \tau_3'(c) > 0 \text{ az } I \text{ intervallumon és } \lim_{c \rightarrow 1} \tau_3(c) = 0.$$

A határértékre vonatkozó állítás nyilván teljesül, az egyenlőtlenség pedig az alábbi.

$$\begin{aligned} & 4(24c^7 + 21c^6 + 39c^5 - 60c^4 + 44c^3 - 36c^2 + 13c + 3) \log c \\ & + 2 \frac{(c-1)}{c} (10c^7 - 26c^6 - c^5 - 25c^4 - 3c^3 - 27c^2 - 18c - 6) \leq 0 \end{aligned} \quad (4.112)$$

Először igazoljuk, hogy a  $\log c$  kifejezés előtt álló tényező szigorúan pozitív az  $I$  intervallumon.

Az  $f_1(c) = c(100c^2 - 144c + 52)$  függvény pozitív  $I$ -n. Az  $f_2(c) = 7c^6 + 12c^5 - 14c^4 + 2$  függvénynek két szélsőértéke van: egy lokális maximuma a  $c_1 = 0$  helyen és egy lokális minimuma a  $c_2 = \frac{-15 + \sqrt{813}}{21}$  helyen. Az  $f_2(c_1), f_2(c_2) > 0$  egyenlőtlenség miatt  $f_2(c)$  is pozitív. A fenti (4.112) egyenletben  $\log c$  együtthatója pedig

$$96c^7 + 6f_2(c) + f_1(c) + 36c^3(1 - c), \quad (4.113)$$

ami szigorúan pozitív  $I$ -n. Ennek a figyelembe vételével átrendezhetjük a tagokat a fenti egyenlőtlenségben, és a következő adódik.

$$q(c) = \log(c) + \frac{2(c-1)(10c^7 - 26c^6 - c^5 - 25c^4 - 3c^3 - 27c^2 - 18c - 6)}{c(96c^7 + 84c^6 + 156c^5 - 240c^4 + 176c^3 - 144c^2 + 52c + 12)} \geq 0 \quad (4.114)$$

Ezt két lépésben bizonyítjuk be.

1. Ha  $0 < c < \frac{1}{2}$ , akkor csökkentjük a  $q(c)$  függvényt és megmutatjuk, hogy még a

$$q^*(c) = \log(c) + \frac{3(c-1)(2c^7 - 1)}{c(24c^7 + 24c^6 + 44c^5 - 60c^4 + 44c^3 - 36c^2 + 13c + 13)} \geq 0 \quad (4.115)$$

egyenlőtlenség is teljesül. Mivel  $q^*(\frac{1}{2}) > 0$ , ezért elég megmutatni, hogy  $\frac{dq^*(c)}{dc} < 0$  teljesül, ha  $0 < c < \frac{1}{2}$ . Ez pedig következik az

$$\begin{aligned} \frac{dq^*(c)}{dc} &= - \frac{1}{c^2(24c^7 + 24c^6 + 44c^5 - 60c^4 + 44c^3 - 36c^2 + 13c + 3)^2} \\ &\cdot \left( c^{12}(-576c^3 - 1440c^2 - 3216c + 2112) + c^{10}(-2944c + 1472) \right. \\ &\quad + c^8(5296c^2 - 7700c + 2640) + c^6(4800c^2 - 7292c + 2508) \\ &\quad \left. + c^3(1800c^3 - 568c^2 - 624c + 241) + (550c^3 - 441c^2 + 69c + 9) \right) \end{aligned} \quad (4.116)$$

egyenlőtlenségből, ahol  $\frac{dq^*(c)}{dc}$  hat negatív értékeket felvevő függvény összege.



2. Ha  $\frac{1}{2} < c < 1$ , akkor mivel  $q(1) = 0$ , ezért elég megmutatni, hogy ha  $\frac{1}{2} < c < 1$ , akkor  $q'(c) < 0$ . Tekintsük a következő függvényt.

$$\zeta(c) = -c^2(24c^7 + 24c^6 + 44c^5 - 60c^4 + 44c^3 - 36c^2 + 13c + 3)^2 \cdot q'(c) \quad (4.117)$$

A cél a  $\zeta(c)$  pozitivitásának a bizonyítása. Mivel

$$\zeta\left(\frac{1}{2}\right), \zeta'\left(\frac{1}{2}\right), \zeta^{(2)}\left(\frac{1}{2}\right), \dots, \zeta^{(9)}\left(\frac{1}{2}\right) > 0 \quad (4.118)$$

teljesül, ezért elég megmutatni, hogy  $\zeta^{(10)}$  pozitív, ha  $\frac{1}{2} < c < 1$ . Ez pedig következik a

$$\zeta^{(10)}(c) = 39916800(104832c^4 + 110292c^3 + 50688c^2 + 3963c + 1015) \quad (4.119)$$

egyenlőségből.

□

#### 4.4. Tétel. A skalárgörbületben szereplő $\beta_{1,k}$ függvény $m$ -monoton.

*Bizonyítás.* A (4.89) és a (4.91) értelmezés alapján

$$2\varphi(a, a, \lambda_k) + \varphi(a, \lambda_k, a) = \frac{1}{\lambda_k} \cdot \tau(a/\lambda_k), \quad (4.120)$$

ahol

$$\tau(c) = -\frac{2c-3}{2c \log^2 c} + \frac{1}{c(1-c) \log c} + \frac{c}{2(1-c)^2}. \quad (4.121)$$

Ezek alapján a  $\beta_{1,k}$  függvény a

$$\beta_{1,k}(a-x, b+x) = \frac{1}{\lambda_k} \cdot \left( \tau\left(\frac{a-x}{\lambda_k}\right) + \tau\left(\frac{b+x}{\lambda_k}\right) \right) \quad (4.122)$$

alakban írható fel. Ez azt jelenti, hogy az  $x \mapsto \beta_{1,k}(a-x, b+x)$  függvény monoton növekvő, ha  $\tilde{x} \mapsto \tau(\tilde{a} - \tilde{x}) + \tau(\tilde{b} + \tilde{x})$  monoton növekvő az  $\tilde{x} \in [0, \frac{\tilde{a} + \tilde{b}}{2}]$  intervallumon, ahol  $\tilde{a} = a/\lambda_k$ ,  $\tilde{b} = b/\lambda_k$  és  $\tilde{x} = x/\lambda_k$ . Az

$$-\tau'(\tilde{a} - \tilde{x}) + \tau'(\tilde{b} + \tilde{x}) > 0 \quad (4.123)$$

egyenlőtlenség azt jelenti, hogy  $\tau'$  monoton csökkenő. Ezt úgy bizonyítjuk, hogy belátjuk a

$$\tau''(\tilde{x}) < 0 \quad (4.124)$$

egyenlőtlenséget.

Tekintsük az alábbi függvényeket.

$$\tau_1(c) = \frac{3-c}{4c \log^2 c} - \frac{1}{2c(c-1) \log c}, \quad \tau_2(c) = \frac{3-3c}{2c \log^2 c} - \frac{1}{c(c-1) \log c} + \frac{c}{(1-c)^2} \quad (4.125)$$

Mivel  $\tau(c) = \tau_1(c) + (1/2)\tau_2(c)$ , ezért a (4.124) egyenlőtlenséget két lépésben bizonyítjuk be, igazoljuk a  $\tau_1$  és a  $\tau_2$  függvényekről, hogy konkávak.

A  $\tau_1$  esetén tekintsük a

$$\tau^*(c) = -2c^3(c-1)^3 \log^4 c \cdot \tau_1''(c) \quad (4.126)$$

segédfüggvényt, ami más alakban

$$\begin{aligned} \tau^*(c) = & (6c^2 - 6c + 2) \log^3 c - (c-1)(3c-2)(c-3) \log^2 c + \\ & + (c-1)^2(c^2 - 10c + 11) \log c + 3(c-1)^3(c-3) . \end{aligned} \quad (4.127)$$

A  $\tau_1$  függvény akkor konkáv, ha  $\tau^*(c)$  negatív  $c \in I$  esetén és pozitív a  $c > 1$  esetben. Mivel  $\lim_{c \rightarrow 1} \tau^*(c) = 0$ , ezért elég igazolni, hogy  $\tau^{*(1)}(c) > 0$ . Az előző gondolatmenet alapján: mivel  $\lim_{c \rightarrow 1} \tau^{*(1)}(c) = 0$ , ezért elég igazolni, hogy  $\tau^{*(2)}(c)$  negatív  $c \in I$  esetén és pozitív, ha  $c > 1$ . Megint alkalmazva az előbbi módszert: mivel  $\lim_{c \rightarrow 1} \tau^{*(2)}(c) = 0$ , ezért elég igazolni, hogy  $\tau^{*(3)}(c) > 0$ , vagy a vele ekvivalens

$$\begin{aligned} \rho(c) = & c^3 \cdot \tau^{*(3)}(c) = (-18c^3 + 36c^2 - 18c + 12) \log^2 c \\ & + (24c^4 - 138c^3 + 164c^2 + 34c - 12) \log c + (98c^4 - 276c^3 - 184c^2 - 4c - 2) > 0 \end{aligned} \quad (4.128)$$

egyenlőtlenséget.

Az előző gondolatmenetet használjuk a  $\rho$  függvény pozitivitásának a bizonyításához. Mivel

$$\lim_{c \rightarrow 1} \rho^{(0)}(c) = \lim_{c \rightarrow 1} \rho^{(1)}(c) = \lim_{c \rightarrow 1} \rho^{(2)}(c) = \lim_{c \rightarrow 1} \rho^{(3)}(c) = 0 \quad (4.129)$$

ezért elég a  $\rho^{(4)} > 0$  egyenlőtlenséget igazolni. A

$$\rho^{(4)}(c) = \frac{1}{c^4} (72(8c^4 - 3c^3 - 2c^2 + c - 2) \log c + 8(444c^4 - 153c^3 - 50c^2 + 4c + 42)) , \quad (4.130)$$

kifejezés pozitivitása a log függvény sorfejtéséből következik. Mutatunk egy  $\rho^{(4)}$  függvénynél kisebb kifejezést, mely még mindig pozitív.

1. Ha  $c > 1$ , akkor a log függvény sorfejtése alapján ismert, hogy a  $c > 1$  esetben

$$\log c > 2 \frac{c-1}{c+1} \quad (4.131)$$

teljesül. Ezért a  $2 \frac{c-1}{c+1}$  tag  $\log c$  helyére való helyettesítése után a  $\rho^{(4)}$  kifejezésben, a következőt kapjuk

$$\frac{1}{c^4(1+c)} \cdot (401c^5 + 185c^3(c^2 - 1) + 93c^4 + 8c(c-1) + 78) . \quad (4.132)$$

Ez a függvény azonban biztosan pozitív a  $c > 1$  esetben.

2. Ha  $0 < c < 1$ , akkor az  $f_1(c) = 8c^4 \log c + 1$ ,  $f_2(c) = -(3c^3 + 2c^2 - c + 2) \log c$  és  $f_3(c) = (444c^4 - 153c^3 - 50c^2 + 4c + 12)$  függvények pozitívak a  $]0, 1]$  intervallumon. Ezek alapján

$$\rho^{(4)}(c) = \frac{1}{c^4} \cdot (72f_1(c) + 72f_2(c) + 8f_3(c) + 312) > 0 \quad \text{ha } 0 < c < 1 \quad (4.133)$$

teljesül

Eddig a (4.125) képlettel értelmezett  $\tau_1$  függvényről bizonyítottuk be, hogy konkáv. A (4.125) képlettel definiált  $\tau_2$  esetén tekintsük az alábbi segédfüggvényt.

$$\tau^*(c) = c^3(c-1)^4 \log^4 c \cdot \tau_2''(c) , \quad (4.134)$$

ami más alakban

$$\begin{aligned} \tau^*(c) = & 2c^3(c+2) \log^4 c - 2(c-1)(3c^2 - 3c + 1) \log^3 c + (3c-2)(c-3)(c-1)^2 \log^2 c \\ & - (3c^2 - 12c + 11)(c-1)^3 \log c - 9(c-1)^5 . \end{aligned} \quad (4.135)$$

A  $\tau_2$  függvény konkáv, ha a  $\tau^*$  függvény pozitív. A

$$\lim_{c \rightarrow 1} \tau^*(c) = \lim_{c \rightarrow 1} \tau^{*(1)}(c) = \dots = \lim_{c \rightarrow 1} \tau^{*(5)}(c) = 0 \quad (4.136)$$

határértékek könnyen ellenőrizhetők, ezért elég igazolni, hogy a

$$\begin{aligned} \rho(c) = \frac{1}{c^2} \tau^{*(5)}(c) = & 48(2c-1) \log^3 c + \frac{6}{c^3} (100c^4 - 11c^3 + 12c^2 + 12c + 12) \log^2 c \\ & - \frac{2}{c^3} (90c^5 - 246c^4 - 216c^3 - 62c^2 - 9c + 78) \log c \\ & - \frac{1}{c^3} (951c^5 - 642c^4 - 403c^3 - 36c^2 + 100c - 42) \end{aligned} \quad (4.137)$$

függvény pozitív  $0 < c < 1$  esetén és negatív, ha  $1 < c$ .

Az előző módszer ismételt alkalmazásával kapjuk, hogy a

$$\lim_{c \rightarrow 1} \rho(c) = \lim_{c \rightarrow 1} \rho'(c) = \lim_{c \rightarrow 1} \rho''(c) = 0 \quad (4.138)$$

határértékek miatt elég igazolni a

$$\begin{aligned} \frac{c^5}{2} \cdot \rho^{(2)}(c) &= (144c^4 + 72c^3 + 72c^2 + 216c + 432) \log^2 c + (-180c^5 + 888c^4 - 78c^3 \\ &\quad - 92c^2 - 306c - 1440) \log c - (1221c^5 - 1056c^4 + 282c^3 + 150c^2 + 273c - 870) \end{aligned} \quad (4.139)$$

függvényről, hogy pozitív, ha  $0 < c < 1$  és negatív, ha  $1 < c$ . Mivel

$$\lim_{c \rightarrow 1} \frac{c^5}{2} \cdot \rho^{(2)}(c) = 0, \quad (4.140)$$

ezért igazoljuk, hogy a

$$\eta(c) = -\frac{1}{576c^3 + 216c^2 + 144c + 216} \cdot \frac{d}{dc} \left( \frac{c^5}{2} \cdot \rho^{(2)}(c) \right) \quad (4.141)$$

függvény, ami egyszerűsítések után a

$$\begin{aligned} \eta(c) &= -\log^2 c + \frac{2(450c^5 - 1920c^4 + 45c^3 + 20c^2 - 63c - 432)}{c(576c^3 + 216c^2 + 144c + 216)} \cdot \log c \\ &\quad + \frac{6285c^5 - 5112c^4 + 924c^3 + 392c^2 + 579c + 1440}{c(576c^3 + 216c^2 + 144c + 216)} \end{aligned} \quad (4.142)$$

alakra hozható, szigorúan pozitív.

Ezt két részletben mutatjuk meg, a log függvény sorfejtése alapján.

1. Ha  $c > 1$ , akkor felhasználjuk, hogy  $c > 1$  esetén  $\sqrt{c} > \log c$  teljesül. Ha  $-c$ -t helyettesítünk  $-\log^2 c$  helyére az  $\eta$  függvényben, akkor csökkentjük az értékét. Az így kapott függvényt tovább csökkentjük és megszorozzuk a  $1152c(11c^3 + 2c + 3)$  taggal, így a

$$\begin{aligned} &-(91905c^4 - 9081c^3 - 9064c^2 + 2016c + 13824) \log c + \\ &+ \frac{(33c^3 + 6c + 9)(1855c^3 - 1776c^2 - 72)}{c} \end{aligned} \quad (4.143)$$

kifejezést kapjuk, és ennek a pozitívását igazoljuk.

Mivel ebben a kifejezésben a  $\log c$  együtthatója szigorúan negatív  $c > 1$  esetén, ezért elég igazolni, hogy a

$$h(c) = -\log c + \frac{3(11c^3 + 2c + 3)(1855c^3 - 1776c^2 - 72)}{c(91905c^4 - 9081c^3 - 9064c^2 + 2016c + 13824)} \quad (4.144)$$

függvény pozitív, ha  $c > 1$ . Egyszerű számolással adódik, hogy  $h(1) = \frac{3}{800} > 0$  és

$$h'(c) = \frac{864(87213719c^4 + 366253c^3 + 339642c^2 + 218160c + 10368) + c^4 q(c)}{c^2(91905c^4 - 9081c^3 - 9064c^2 + 2016c + 13824)^2}, \quad (4.145)$$

ahol  $q(c)$  a  $c$  polinomja. A  $h'$  függvény pozitivitása az  $1 < c$  esetben, következik a

$$q(1), q'(1), q^{(2)}(1), q^{(3)}(1) > 0 \quad (4.146)$$

egyenlőtlenségekből, valamint a

$$q^{(4)}(c) = 2376(852418875c^2 - 482743225c - 4909728) > 0 \quad \text{ha } 1 < c \quad (4.147)$$

relációkból.

2. Ha  $0 < c < 1$ , akkor, mivel a (4.141) képlettel definiált  $\eta$  függvényre  $\eta(1) > 0$  teljesül, ezért elég igazolni a

$$36c^2(8c^3 + 3c^2 + 2c + 3) \cdot \frac{d\eta(c)}{dc} < 0 \quad (4.148)$$

egyenlőtlenséget. Az előbbi egyenlőtlenség egyszerűsített alakja

$$\begin{aligned} & (3600c^8 - 1908c^7 - 6876c^6 - 5552c^5 - 20058c^4 + 12888c^3 + \\ & + 3210c^2 + 1080c + 1296) \cdot \log c + (28740c^8 + 4845c^7 + 2991c^6 + \\ & + 22155c^5 - 35730c^4 - 25475c^3 - 7833c^2 - 3933c - 3456). \end{aligned} \quad (4.149)$$

Igazoljuk, hogy növelve az előbbi kifejezést, és megszorozva  $c^3$ -bel

$$\begin{aligned} & -2c(954c^3 + 3438c^2 + 2776c + 10029) \cdot \log c + (28740c^5 + 4845c^4 + 2991c^3 \\ & + 22155c^2 - 35730c - 25475) \end{aligned} \quad (4.150)$$

még mindig negatív marad a  $0 < c < 1$  esetben. A kifejezésben szereplő  $\log c$  együtthatója negatív  $0 < c < 1$  esetén, ezért elég igazolni, hogy a

$$\eta^*(c) = \log c - \frac{28740c^5 + 4845c^4 + 2991c^3 + 22155c^2 - 35730c - 25475}{2c(954c^3 + 3438c^2 + 2776c + 10029)} \quad (4.151)$$

függvény pozitív, ha  $0 < c < 1$ . Az  $\eta^*(1) > 0$  egyenlőtlenség miatt elég megmutatni, hogy

$$-2c^2(954c^3 + 3438c^2 + 2776c + 10029)^2 \cdot \frac{d\eta^*(c)}{dc} \quad (4.152)$$

pozitív. Ezt egyszerűsítve a

$$(472766109c^2 - 59724482c + 25548875) + a_3c^3 + a_4c^4 + \dots + a_8c^8 \quad (4.153)$$

alakot kapjuk, ahol  $a_3, a_4, \dots, a_8 > 0$ , és az első tag pozitív, ha  $0 < c < 1$ .

Ezzel igazoltuk, hogy a  $\tau_1$  és a  $\tau_2$  függvény konkáv.  $\square$

A (4.100) skalárgörbületi formában szereplő  $\beta_{2,k}$  függvény m-monotonitása azt jelenti, hogy minden  $1 > a > b$  pozitív szám és  $1 > \lambda_k > 0$  sajátérték esetén,  $y \mapsto \beta_{2,k}(a - y, b + y)$  szigorúan monoton csökkenő függvénye  $y$ -nak az  $y \in [0, \frac{a-b}{2}]$  intervallumon.

Bevezetve a  $c = \frac{(a+b)}{2\lambda_k}$  jelölést a  $\beta_{2,k}$  függvény m-monotonitása az  $x \mapsto \beta_{2,k}(c+x, c-x)$  függvény monoton csökkenését jelenti az  $x \in [0, c]$  esetben.

Definiáljuk az alábbi függvényeket

$$\phi_1(u) = \frac{1}{u \log u} + \frac{1}{1-u}, \quad \phi_2(u) = \frac{1}{\log u} + \frac{1}{1-u}. \quad (4.154)$$

A (4.89) egyenlet alapján kapjuk, hogy

$$\lambda_k \cdot \beta_{2,k}(a, b) = 3w_\beta(x, c) + 2q_\beta(x, c) + 2r_\beta(x, c), \quad (4.155)$$

ahol

$$w_\beta(x, c) = \frac{m(c-x, c+x, 1)^2}{m(c-x, c+x)m(c-x, 1)m(c+x, 1)}, \quad (4.156)$$

$$r_\beta(x, c) = -\frac{m(c-x, 1, 1)m(c+x, 1, 1)}{m(c-x, 1)m(c+x, 1)},$$

$$q_\beta(x, c) = \frac{-m(c-x, c-x, c+x)m(c-x, c-x, 1)}{m(c-x, c-x)m(c-x, 1)m(c-x, c+x)} - \frac{m(c+x, c+x, c-x)m(c+x, c+x, 1)}{m(c+x, c+x)m(c+x, 1)m(c-x, c+x)}.$$

Ezen függvények explicit alakja a következő:

$$w_\beta(x, c) = \frac{t_\beta(x, c) + t_\beta(-x, c)}{2}, \quad \text{ahol} \quad t_\beta(x, c) = \frac{\frac{\log(c+x)}{\log(c-x)} \cdot \frac{c-x-1}{c+x-1} - 1}{x(\log(x+c) - \log(x-c))}, \quad (4.157)$$

$$q_\beta(x, c) = q_{1,\beta}(x, c) + q_{2,\beta}(x, c), \quad \text{ahol} \quad \begin{cases} q_{1,\beta}(x, c) = \frac{\phi_1(c+x) - \phi_1(c-x)}{\log(x+c) - \log(x-c)}, \\ q_{2,\beta}(x, c) = \frac{\phi_2(c+x)}{x} + \frac{\phi_2(c-x)}{-x}, \end{cases}$$

$$r_\beta(x, c) = -\phi_2(c-x)\phi_2(c+x).$$

A  $\beta_{2,k}$  függvény  $m$ -monotonitása következik a

$$\frac{d}{dx}(3w_\beta(x, c) + 2q_\beta(x, c) + 2r_\beta(x, c)) < 0 \quad (4.158)$$

egyenlőtlenségből. Ezt az egyenlőtlenséget elég nehéz igazolni, azonban fel lehet bontani négy tagra, melyekről a numerikus szimulációk azt mutatják, hogy külön-külön negatívak.

**4.5. Tétel.** *Ha minden pozitív  $c$  paraméter mellett az  $x \in ]0, c[$  intervallumon teljesülnek az*

$$1: \frac{d}{dx}(w_\beta(x, c) + q_\beta(x, c)) < 0 \quad 2: \frac{d}{dx}(2w_\beta(x, c) + q_{1,\beta}(x, c) + r_\beta(x, c)) < 0 \quad (4.159)$$

$$3: \frac{d}{dx}(q_{2,\beta}(x, c)) < 0 \quad 4: \frac{d}{dx}(r_\beta(x, c)) < 0$$

*egyenlőtlenségek, akkor a skalárgörbületben szereplő  $\beta_{2,k}$  függvény  $m$ -monoton.*

**4.6. Tétel.** *Az előző tételben szereplő 4. feltétel teljesül.*

*Bizonyítás.* A 4. feltétel

$$\frac{d}{dx}(r_\beta(x, c)) < 0, \quad (4.160)$$

amivel ekvivalens az

$$\frac{\phi_2'(c+x)}{\phi_2(c+x)} > \frac{\phi_2'(c-x)}{\phi_2(c-x)} \quad (4.161)$$

egyenlőtlenség, ahol a  $\phi_2$  függvényt a (4.154) egyenlet definiálja. Ezért elég a

$$\left(\frac{\phi_2'(u)}{\phi_2(u)}\right)' > 0 \quad (4.162)$$

feltétel teljesülését, vagy a vele ekvivalens

$$\rho(u) = \phi_2(u)\phi_2''(u) - (\phi_2'(u))^2 > 0 \quad (4.163)$$

egyenlőtlenséget igazolni.

Megmutatjuk, hogy a  $\rho_*(u) = (1-u)^4 \log^4 u \cdot \rho(u)$  függvény pozitív. A

$$\rho_*(1) = \rho_*^{(1)}(1) = \rho_*^{(2)}(u) = \rho_*^{(3)}(1) = 0 \quad (4.164)$$

egyenlőségek miatt elég igazolni, hogy a  $\tau(u) = u^4 \cdot \rho_*^{(3)}(u)$  függvény, ami explicit alakban

$$\begin{aligned} \tau(u) = & 2(-3u^3 - 6u^2 + u - 6) \log^2 u - 2(1-u)(2u^3 - 9u^2 - 3u + 2) \log u + \\ & + 2(1-u)^2(6u^2 + 5u + 11), \end{aligned} \quad (4.165)$$

$$(4.166)$$

negatív  $0 < u < 1$  esetén, és pozitív, ha  $1 < u$ . A

$$\tau(1) = \tau^{(1)}(1) = \rho^{(2)}(1) = \tau^{(3)}(1) = \tau^{(4)}(1) = 0 \quad (4.167)$$

egyenlőségek miatt elég igazolni, hogy a

$$\tau_*(u) = u^4 \cdot \tau^{(4)}(u) \quad (4.168)$$

függvény, mely explicit alakban

$$\tau_*(u) = (96u^4 - 72u^3 + 48u^2 + 8u + 144) \log u - 8(1-u)(61u^3 + 28u^2 + 28u + 30), \quad (4.169)$$

monoton növény, vagy bebizonyítani a vele ekvivalens

$$\tau_*^{(1)}(u) = (384u^3 - 216u^2 + 96u + 8) \log u + \frac{8}{u}(256u^4 - 108u^3 + 6u^2 + 3u + 18) > 0 \quad (4.170)$$

egyenlőtlenséget. Pozitív  $u$  esetén a  $\log u$  tag együtthatója pozitív. A fenti egyenlőtlenséget elosztva  $\log u$  együtthatójával kapjuk a

$$d(u) = \log u + \frac{16}{3} + \frac{18}{u} + \frac{1}{3} \cdot \frac{2484u^2 - 1284u + 655}{48u^3 - 27u^2 + 12u + 1} > 0 \quad (4.171)$$

egyenlőtlenséget, melyet három lépésben igazolunk.



1. Ha  $0 < u < \frac{1}{2}$ , akkor a  $d(\frac{1}{2}) > 0$  egyenlőtlenség miatt elég megmutatni, hogy a

$$d_*(u) = u^2(48u^3 - 27u^2 + 12u + 1)^2 \cdot d'(u) \quad (4.172)$$

kifejezés negatív, ha  $0 < u < \frac{1}{2}$ . Mivel  $d_*(0) < 0$ , ezért a  $d'_* < 0$  egyenlőtlenséget kell csak igazolni. Ez azonban következik a

$$d'_*(u) = -u^5(25920 - 16128u) - u^2(-37245u^2 + 5400u^3 + 5000) - (5260u^2 - 2904u + 431) \quad (4.173)$$

egyenletből, ahol  $d'_*$  három negatív függvény összegeként áll elő.

2. Ha  $\frac{1}{2} < u < 1$ , akkor könnyen igazolható, hogy a

$$c(u) = \frac{1}{3} \cdot \frac{2484u^2 - 1284u + 655}{48u^3 - 27u^2 + 12u + 1} \quad (4.174)$$

függvény konkáv, ezért

$$c(u) > \frac{c(1) - c(1/2)}{1 - 1/2} \cdot (u - 1/2) + c(1/2), \quad \text{ha } \frac{1}{2} < u < 1 \quad (4.175)$$

teljesül. A fenti egyenlőtlenség jobb oldalát behelyettesítve a (4.171) egyenlőtlenségbe

$$\log(u) - \frac{113323}{2550} + \frac{18}{u} + \frac{13283}{425}u > 0 \quad (4.176)$$

adódik. Ennek az egyenlőtlenségnek a bal oldalát tovább csökkentjük, és igazoljuk a

$$c_*(u) = \log u - 45 + \frac{18}{u} + 30u > 0 \quad \text{ha } \frac{1}{2} < u < 1 \quad (4.177)$$

egyenlőtlenséget. A  $c_*(u)$  függvénynek egy szélsőértéke van az  $u_0 = \frac{\sqrt{2161}-1}{60}$  helyen, ami lokális minimum, és  $c_*(u_0) > 0$  miatt teljesül a fenti egyenlőtlenség.

3. Ha  $1 < u$ , akkor, mivel a (4.171) képlettel definiált  $d$  függvényre  $d(1) > 0$  teljesül, ezért elég a

$$c(u) = u^2(48u^3 - 27u^2 + 12u + 1)^2 \cdot d'(u) \quad (4.178)$$

függvény pozitivitását igazolni az  $1 < u$  esetben. A

$$c(1), c'(1), c^{(2)}(1), c^{(3)}(1), c^{(4)}(1) > 0 \quad (4.179)$$

egyenlőtlenségek miatt elég csak  $c^{(5)}(u) > 0$  igazolása az  $1 < u$  esetben. Ez azonban következik a

$$c^{(5)}(u) = 5806080u^2 - 3110400u + 893880 \quad (4.180)$$

egyenlőségéből.

□

A 4.5. tétel 1., 2. és 3. feltételei remélhetőleg hasonló módon igazolhatóak. A végezett numerikus szimulációk megerősítik teljesülésüket: A  $[0, 10^3]$  paraméterteréből egyenletesen választva  $10^4$  számú  $c$  értéket, és minden  $c$  érték esetén egyenletesen választva  $10^4$  számú  $x$ -et a  $[0, c]$  intervallumból az említett három feltétel teljesül.

A skalárgörbület (4.100) képletében szereplő  $\gamma_{kl}$  tagot a (4.89), (4.91) egyenlőségek segítségével határozhatjuk meg

$$\begin{aligned} \gamma_{kl}(a, b) = \frac{1}{\lambda_l} & \left( 3w(\tilde{a}, \tilde{\lambda}_k) + 3w(\tilde{b}, \tilde{\lambda}_k) + q(\tilde{a}, \tilde{\lambda}_k) + q(\tilde{b}, \tilde{\lambda}_k) + \right. \\ & \left. + r(\tilde{a}, \tilde{\lambda}_k) + r(\tilde{b}, \tilde{\lambda}_k) + d(\tilde{a}, \tilde{\lambda}_k) + d(\tilde{b}, \tilde{\lambda}_k) \right), \end{aligned} \quad (4.181)$$

ahol  $\tilde{a} = \frac{a}{\lambda_l}$ ,  $\tilde{b} = \frac{b}{\lambda_l}$ ,  $\tilde{\lambda}_k = \frac{\lambda_k}{\lambda_l}$  és

$$w(x, c) = \frac{1}{(c-1)\log c} \left( \frac{\log c}{\log x - \log c} + \frac{1-c}{x-c} \right) \left( \frac{\log c}{\log x} + \frac{c-1}{1-x} \right), \quad (4.182)$$

$$q(x, c) = \phi_1(c)\phi_2\left(\frac{x}{c}\right),$$

$$r(x, c) = \frac{-1}{x}\phi_2\left(\frac{c}{x}\right)\phi_2\left(\frac{1}{x}\right),$$

$$d(x, c) = \phi_2(c)\phi_2(x).$$

(A  $\phi_1$  és  $\phi_2$  függvényt a (4.154) egyenlet definiálta.) A  $\gamma_{kl}$  függvény  $m$ -monotonitása azt jelenti, hogy  $x \mapsto \gamma_{kl}(a-x, b+x)$  monoton növvő függvénye  $x$ -nek az  $x \in [0, \frac{a-b}{2}]$  esetben. Ez következik a

$$\frac{d^2}{dx^2} (3w(x, c) + q(x, c) + d(x, c) + r(x, c)) < 0 \quad (4.183)$$

egyenlőtlenségből. Ezt nehéz az eddigiekből igazolni, de az egyenlőtlenség bal oldala felbontható két, látszólag negatív tag összegére.

**4.7. Tétel.** *Ha minden pozitív  $c$  paraméter esetén minden pozitív  $x$ -re a*

$$1: \quad \frac{d^2}{dx^2} (2w(x, c) + d(x, c)) < 0 \quad 2: \quad \frac{d^2}{dx^2} (w(x, c) + q(x, c) + r(x, c)) < 0 \quad (4.184)$$

*feltételek teljesülnek, akkor a (4.100) skalárgörbületben szereplő  $\gamma_{k,l}$  függvény  $m$ -monoton.*

Az előbb említett két feltételt a numerikus számítások megerősítik a  $[0, 10^7]$  paraméterterén.

Érdeemes megjegyezni, hogy a skalárgörbületben szereplő tagok (4.100) egyenletben szereplő csoportosítása nem az egyetlen lehetőség, azonban a numerikus és az elméleti vizsgálatok ezt helyezik előtérbe. Az (4.100) egyenletben szereplő tagokról többet is állíthatunk az itt leírtakon kívül. Ezen állítások bizonyítását mellőzzük, mert azok az előbbiekhöz hasonló, de komplikáltabb számolásokat igényelnek. (A fentebb leírt bizonyítások egyszerűsített, rövidített formái az alábbi állítások bizonyításának.)

1.  $\alpha(a, b)$ :

Adott  $a > b > 0$  paraméterek esetén az

$$x \mapsto \varphi(a - x, a - x, b + x) + \varphi(b + x, b + x, a - x), \quad (4.185)$$

$$x \mapsto \varphi(a - x, b + x, a - x) + \varphi(b + x, a - x, b + x)$$

függvények szigorúan monotonak az  $x \in [0, \frac{a-b}{2}]$  esetben. (Ebből következik a 4.3. tétel.)

2.  $\beta_{1,k}(a, b)$ :

Adott  $1 > a > b > 0$  és  $1 > \lambda_k > 0$  paraméterek esetén az

$$x \mapsto \varphi(a - x, a - x, \lambda_k) + \varphi(b + x, b + x, \lambda_k), \quad (4.186)$$

$$x \mapsto \varphi(a - x, a - x, \lambda_k) + \varphi(b + x, b + x, \lambda_k) + \\ + \varphi(a - x, \lambda_k, a - x) + \varphi(b + x, \lambda_k, b + x)$$

függvények szigorúan monoton növekednek, azonban az

$$x \mapsto \varphi(a - x, \lambda_k, a - x) + \varphi(b + x, \lambda_k, b + x) \quad (4.187)$$

függvény nem monoton növekszik az  $x \in [0, \frac{a-b}{2}]$  tartományon.

3.  $\beta_{2,k}(a, b)$ :

Adott  $1 > a > b > 0$  és  $1 > \lambda_k > 0$  paraméterek esetén az

$$x \mapsto \varphi(a - x, b + x, \lambda_k) + \varphi(b + x, a - x, \lambda_k), \quad (4.188)$$

$$x \mapsto \varphi(a - x, b + x, \lambda_k) + \varphi(b + x, a - x, \lambda_k) + \\ + \varphi(a - x, \lambda_k, b + x) + \varphi(b + x, \lambda_k, a - x)$$

függvényekről a numerikus szimuláció azt mutatja, hogy szigorúan monoton növekvő, azonban az

$$x \mapsto \varphi(a - x, \lambda_k, b + x) + \varphi(b + x, \lambda_k, a - x) \quad (4.189)$$

függvény nem monoton növekvő az  $x \in [0, \frac{a-b}{2}]$  esetben.

4.  $\gamma_{k,l}(a, b)$ :

Adott  $1 > a > b > 0$  és  $1 > \lambda_k, \lambda_l > 0$  paraméterek esetén az

$$x \mapsto \varphi(\lambda_k, a - x, \lambda_l) + \varphi(\lambda_k, b + x, \lambda_l) \quad (4.190)$$

függvényről a numerikus szimuláció azt mutatja, hogy szigorúan monoton növekvő, azonban az

$$x \mapsto \varphi(a - x, \lambda_k, \lambda_l) + \varphi(b + x, \lambda_k, \lambda_l) \quad (4.191)$$

függvény nem monoton növekvő az  $x \in [0, \frac{a-b}{2}]$  esetben.

Ha a (4.91) képlettel értelmezett  $\varphi$  függvénynek az alábbi módon szimmetrizált változatát tekintjük

$$\begin{aligned} \gamma_{kl}^*(a, b) = & \varphi(a, \lambda_k, \lambda_l) + \varphi(b, \lambda_k, \lambda_l) + \varphi(\lambda_k, a, \lambda_l) + \\ & + \varphi(\lambda_k, b, \lambda_l) + \varphi(\lambda_l, \lambda_k, a) + \varphi(\lambda_l, \lambda_k, b), \end{aligned} \quad (4.192)$$

akkor az  $x \mapsto \gamma_{k,l}^*(a - x, b + x)$  függvény (numerikusan) szigorúan monoton növekvőnek tűnik, azonban ha a  $\lambda_k$  és  $\lambda_l$  sajátértékek hányadosa elég nagy ( $\sim 15000$ ), akkor már az  $x \mapsto \gamma_{k,l}^*(a - x, b + x)$  függvény nem lesz monoton növekvő.

#### 4.2.4. Skalárgörbület monotonitása a valós esetben

A valós sűrűségi mátrixok a komplex sűrűségi mátrixok terének részsokaságát alkotják. Egy részsokaság görbületi tenzora nagyon különbözhet az eredeti sokaság görbületi tenzorától. (Például míg a háromdimenziós euklideszi tér görbületi tenzora azonosan 0, addig a benne lévő kétdimenziós gömbfelszín görbületi tenzora már tartalmaz nullától különböző elemeket.) Petz sejtését át lehet fogalmazni a valós sűrűségi mátrixok terére. Az előzőek alapján érdemes megvizsgálni, hogy a két sejtés (a valós ill. a komplex esetre vonatkozó) között van-e valamilyen kapcsolat.

**4.8. Tétel.** *Ha Petz sejtése igaz a komplex sűrűségi mátrixok terén, akkor a valós sűrűségi mátrixok terén is igaz.*

*Bizonyítás.* Az (4.93), (4.94) egyenletek alapján

$$\text{Scal}_{\mathbb{R}}(D) = \frac{1}{4} \text{Scal}(D) + \frac{1}{4} \sum_{k,l=1}^n v(\lambda_k, \lambda_l), \quad (4.193)$$

ahol a  $v$  függvényt a (4.91) képlet definiálta. A valós esetre vonatkozó skalárgörbület  $m$ -monotonitása az alábbi két állításból következik.

a. Minden  $\lambda_k > \lambda_l$  sajátértékre az  $x \mapsto v(\lambda_k - x, \lambda_l + x) + v(\lambda_l + x, \lambda_k - x)$  függvény monoton növekvő  $x \in [0, \frac{\lambda_k - \lambda_l}{2}]$  esetén.

b. Minden  $\lambda_k > \lambda_l$  és  $\lambda_j$  sajátértékre az  $x \mapsto v(\lambda_k - x, \lambda_j) + v(\lambda_l + x, \lambda_j) + v(\lambda_j, \lambda_k - x) + v(\lambda_j, \lambda_l + x)$  függvény monoton növekvő  $x \in [0, \frac{\lambda_k - \lambda_l}{2}]$  esetén.

Az első állításban szereplő függvény a (4.89), (4.91) egyenletek alapján

$$v(\lambda_k - x, \lambda_l + x) + v(\lambda_l + x, \lambda_k - x) = \frac{1}{\lambda_k + \lambda_l} \left[ (1 + c(x))\zeta(c(x)) + \left( 1 + \frac{1}{c(x)} \zeta\left(\frac{1}{c(x)}\right) \right) \right], \quad (4.194)$$

ahol

$$c(x) = \frac{\lambda_k - x}{\lambda_l + x}, \quad \zeta(c) = \frac{3}{2c \log^2 c} - \frac{2c + 1}{c(c - 1) \log c} + \frac{c + 2}{2(c - 1)^2}. \quad (4.195)$$

Mivel  $c$  csökkenő függvény, ezért elég igazolni, hogy a

$$d(c) = (1 + c)\zeta(c) + \left( 1 + \frac{1}{c} \zeta\left(\frac{1}{c}\right) \right) \quad (4.196)$$

függvény csökkenő a  $c > 1$  esetben. Ez következik a

$$d_*(c) = (c - 1)^3 \log^3 c \cdot d'(c) \quad (4.197)$$

kifejezés negativitásából. Mivel

$$\lim_{c \rightarrow 1} d_*(c) = \lim_{c \rightarrow 1} d'_*(c) = 0, \quad (4.198)$$

ezért elég megmutatni, hogy

$$\tau(c) = \frac{1}{(1 - c)(6c^4 + 18c^2 + 6)} \cdot d''_*(c) < 0 \quad \text{ha } 1 < c. \quad (4.199)$$

A  $\lim_{c \rightarrow 1} \tau(c) = 0$  egyenlőség miatt elég a

$$\begin{aligned} \frac{3c(c^4 + 3c^2 + 1)^2}{(1 - c)^2} \cdot \tau'(c) &= (6c^6 + 22c^5 + 59c^4 + 36c^3 + 59c^2 + 22c + 6) \log c \\ &+ \frac{5}{2}(1 - c^2)(c^4 + 6c^3 + 6c + 1) \end{aligned} \quad (4.200)$$

kifejezésről megmutatni, hogy pozitív. A  $\log c$  tag együtthatója szigorúan pozitív, ezért elég a

$$\tau_*(c) = \log c + \frac{5}{2}(1 - c^2) \cdot \frac{c^4 + 6c^3 + 6c + 1}{59c^2 + 22c + 6} < 0 \quad \text{ha } c > 1 \quad (4.201)$$

egyenlőtlenséget igazolni. Mivel  $\lim_{c \rightarrow 1} \tau_*(c) = 0$ , ezért elég megmutatni, hogy a  $\tau'_*$  függvény pozitív. Ez pedig következik a

$$\tau'_*(c) = \frac{a_0 + a_1c + a_2c^2 + \dots + a_{12}c^{12}}{c(6c^6 + 22c^5 + 59c^4 + 36c^3 + 59c^2 + 22c + 6)} \quad (4.202)$$

egyenletből, ahol  $a_0, \dots, a_{12}$  szigorúan pozitív számok. Ezzel igazoltuk az első állítást.

A második állításban szereplő függvény a (4.89), (4.91) egyenletek alapján a

$$\begin{aligned} v(\lambda_k - x, \lambda_j) + v(\lambda_l + x, \lambda_j) + v(\lambda_j, \lambda_k - x) + v(\lambda_j, \lambda_l + x) = \\ = \frac{1}{\lambda_j} (\zeta(\tilde{\lambda}_k - \tilde{x}) + \zeta(\tilde{\lambda}_l + \tilde{x}) + \rho(\tilde{\lambda}_k - \tilde{x}) + \rho(\tilde{\lambda}_l + \tilde{x})) \end{aligned} \quad (4.203)$$

alakba írható, ahol  $\tilde{\lambda}_k = \frac{\lambda_k}{\lambda_j}$ ,  $\tilde{\lambda}_l = \frac{\lambda_l}{\lambda_j}$ ,  $\tilde{x} = \frac{x}{\lambda_j}$ , a  $\zeta$  függvényt a (4.195) képlet definiálja és

$$\rho(c) = \frac{3}{2 \log^2 c} - \frac{c + 2}{(c - 1) \log c} + \frac{1 + 2c}{2(1 - c)^2}. \quad (4.204)$$

A 4.4. tétel bizonyításában használt elv alapján elég igazolni, hogy a  $\zeta + \rho$  függvény konkáv minden pozitív  $c$  számra. Tekintsük az alábbi segédfüggvényt

$$d(c) = \frac{(c - 1)^4 \log^4 c}{c + 5} \cdot (\zeta''(c) + \rho''(c)). \quad (4.205)$$

A második állítás igazolásához elég megmutatni, hogy  $d(c) < 0$  minden  $c > 0$  esetén. A

$$\lim_{c \rightarrow 1} d(c) = \lim_{c \rightarrow 1} d'(c) = \lim_{c \rightarrow 1} d^{(2)}(c) = \dots = \lim_{c \rightarrow 1} d^{(6)}(c) = 0 \quad (4.206)$$

egyenlőségek miatt ez ekvivalens a

$$\begin{aligned} \tau(c) = d^{(6)}(c) = 1440 \log^3 c - \frac{24}{c^5} (30c^6 - 211c^5 - 198c^4 + 207c^3 + 18c^2 + 54c + 90) \log^2 c \\ + \frac{8}{c^5} (1 - c) (351c^5 + 1306c^4 - 29c^3 + 1120c^2 + 957c + 765) \log c \\ + \frac{4}{c^5} (1 - c)^2 (1249c^4 + 1132c^3 - 744c^2 - 872c - 705) < 0 \quad \text{ha } 0 < c \end{aligned} \quad (4.207)$$

egyenlőtlenséggel. A

$$\lim_{c \rightarrow 1} \tau(c) = \lim_{c \rightarrow 1} \tau'(c) = 0, \quad \lim_{c \rightarrow 1} (c^6 \tau'(c))' < 0 \quad \text{és} \quad \tau(4), \tau'(4), (c^6 \tau'(c))'(4) < 0 \quad (4.208)$$

relációk miatt a (4.207) egyenlőtlenség bizonyításához a  $0 < c < 1$  és  $4 < c$  esetekben elég igazolni a

$$\tau_*(c) = (c^6 \cdot \tau'(c))'' \quad (4.209)$$

függvényről, hogy pozitív, ha  $0 < c < 1$  és negatív ha  $4 < c$ .

A (4.207) egyenlőtlenséget négy lépésben igazoljuk.

1. Ha  $0 < c < \frac{1}{2}$ , akkor csökkentve a  $\tau_*$  függvényt és elosztva  $c^2$ -tel a

$$\begin{aligned} \psi(c) = & -78624 \log^2 c + 96 \frac{4483c^2 + 108}{c^3} \log c - 53008c + 194856 - \\ & - \frac{76944}{c} - \frac{16208}{c^2} + \frac{56520}{c^4} > 0 \end{aligned} \quad (4.210)$$

egyenlőtlenséget kapjuk. A  $\psi(\frac{1}{2}) = 0$  egyenlőség miatt a

$$\begin{aligned} \psi_*(c) = & - \frac{c^4 \psi'(c)}{157248c^3 + 430368c^2 + 31104} \\ = & \log c + \frac{16}{c} (3313c^5 - 31707c^3 - 2026c^2 - 648c + 14130) > 0 \end{aligned} \quad (4.211)$$

egyenlőtlenséget kell igazolni. Mivel  $\psi_*(\frac{1}{2}) > 0$ , ezért elég bizonyítani, hogy

$$p(c) = 3c^2(1638c^3 + 4483c^2 + 324)^2 \psi'_*(c) < 0. \quad (4.212)$$

A  $p(c)$  függvény  $c$ -nek egy polinomja lesz, melyről igazolható, hogy negatív a  $0 < c < \frac{1}{2}$  esetben.

2. Ha  $\frac{1}{2} < c < 1$ , akkor a  $\tau_*$  függvényt (melyet a (4.209) egyenlet definiál) csökkentve a

$$-819c^2 \log^2 c + \frac{4483c^2 + 108}{c} \log c - 553c^3 + 2029c^2 - 800c - 169 + \frac{588}{c^2} > 0 \quad (4.213)$$

egyenlőtlenséget kapjuk. A függvény Taylor-sorfejtése alapján ellenőrizhető a

$$\frac{4483c^2 + 108}{c} \log > a_0 + a_1(c-1/2) + a_2(c-1/2)^2, \quad -819c^2 \log^2 c > b_0 + b_1(c-1/2) \quad (4.214)$$

relációk teljesülése, ahol

$$\begin{aligned} a_0 &= -\frac{4915 \log 2}{2}, & a_1 &= 4915 - 4051 \log 2, & a_2 &= 3187 - 864 \log 2, & (4.215) \\ b_0 &= -\frac{819 \log^2 2}{4}, & b_1 &= 819 \log 2(1 - \log 2). \end{aligned}$$

Igazoljuk, hogy a fenti egyenlőtlenség logaritmust tartalmazó tagjai helyére a Taylor-sorfejtés jobb oldalát helyettesítve, még mindig pozitív függvényt kapunk

$$-\frac{1}{4c^2}(2212c^5 + \alpha_4 c^4 + \alpha_3 c^3 + \alpha_2 c^2 - 2352) > 0, \quad (4.216)$$

ahol  $\alpha_4 = 3456 \log 2 - 20864$ ,  $\alpha_3 = 3276 \log^2 2 + 9472 \log 2 - 3712$  és  $\alpha_2 = -819 \log^2 2 + 4230 \log 2 + 7319$ . Könnyen ellenőrizhető, hogy a zárójelben lévő polinom szigorúan negatív  $\frac{1}{2} < c < 1$  esetén.

3. Ha  $1 < c < 4$ , akkor a (4.125) egyenlettel definiált  $\tau$  függvény negativitását úgy igazoljuk, hogy a logaritmust tartalmazó tagokat nagyobbakra cseréljük, és az így kapott kifejezésről mutatjuk meg, hogy még mindig negatív.

A (4.207) kifejezéssel értelmezett  $\tau$  függvénybe helyettesítsük a  $-2500(c-1)$  tagot a

$$-\frac{24}{c^5}(30c^6 - 211c^5 - 198c^4 + 207c^3 + 18c^2 + 54c + 90) \log^2 c \quad (4.217)$$

kifejezés helyére, a  $-2000(c-1)^2$  tagot a

$$\frac{8}{c^5}(1-c)(351c^5 + 1306c^4 - 29c^3 + 1120c^2 + 957c + 765) \log c \quad (4.218)$$

helyére és a  $\frac{4}{c^5} \cdot (1-c)^2(1249c^4 + 1132c^3)$  tagot

$$\frac{4}{c^5}(1-c)^2(1249c^4 + 1132c^3 - 744c^2 - 872c - 705) \quad (4.219)$$

helyére. A Taylor-sorfejtés segítségével igazolhatjuk, hogy ezzel a módszerrel tényleg növeltük a  $\tau$  függvényt. Az új egyenlőtlenség az alábbi

$$\tau_1(c) = 1440 \log^3 c + \frac{4(c-1)}{c^2} \cdot (500c^3 - 1124c^2 + 117c + 1132) < 0 \quad \text{ha } 1 < c < 4. \quad (4.220)$$

Növeljük a  $\tau_1$  függvényt, és igazoljuk a

$$\tau_2(c) = \tau_1(c) + 2000 \frac{(c-1)(c-2)^2}{c^2} < 0 \quad (4.221)$$



egyenlőtlenséget. Mivel  $\lim_{c \rightarrow 1} \tau_2(c) = 0$  elég igazolni, hogy  $c\tau_2'(c)$  negatív. Ehhez növeljük a  $c\tau_2'(c) < 0$  függvényt és igazoljuk a

$$\tau_3(c) = c\tau_2'(c) + 2000 \left(c - \frac{3}{2}\right)^2 + 8(c-4)^2 - 72 < 0 \quad (4.222)$$

egyenlőtlenséget. A  $\lim_{c \rightarrow 1} \tau_3(c) = 0$  egyenlet miatt elég a

$$\tau_3'(c) = 8640 \log c + \frac{4(-996c^4 + 608c^3 + 2985c - 3472)}{c^2} < 0 \quad \text{ha } 1 < c < 4 \quad (4.223)$$

egyenlőtlenség igazolása.

A  $2 < c < 4$  esetben  $(c-1)$ -et helyettesítve  $\log c$  helyére igazolható a

$$\frac{-996c^4 + 2768c^3 - 2160c^2 + 2985c - 3472}{c^2} \quad \text{ha } 2 < c < 4 \quad (4.224)$$

egyenlőtlenség.

Az  $1 < c < 2$  esetben a (4.223) egyenlőtlenség igazolható a (Taylor-sorfejtésből levezethető)

$$\frac{c-1}{c} + \frac{(c-1)^2}{2c^2} + \frac{(c-1)^3}{3c^3} + \frac{(c-1)^4}{4c^4} + \frac{(c-1)^5}{5c^5} + \frac{(c-1)^6}{c^6} > \log c \quad \text{ha } 1 < c < 2 \quad (4.225)$$

egyenlőtlenség segítségével.

4. Ha  $4 < c$ , akkor növelve a  $\tau_*$  függvényt (amit a (4.209) egyenlet definiál) a

$$(86400c^3 + 59616c + 2592) \log^2 c + \frac{48}{c}(-2985c^5 + 5840c^4 + 3126c^2 + 216) \log c + \frac{8}{c^2}(2184c^6 + 24357c^4 + 204c + 7065) < 0$$

egyenlőtlenséget kapjuk. A  $\log^2 c$  és  $\log c$  tag együtthatói szigorúan pozitívak és  $c-2 > \log^2 c$ ,  $c-2 > \log c$ , ha  $4 < c$ , ezért tovább növelhetjük a fenti egyenlőtlenség bal oldalát, ha  $(c-2)$ -t helyettesítünk  $\log^2 c$  és  $\log c$  helyére. Ekkor a

$$c^4(5970c^3 - 27948c^2 + 30560c - 16855) + (17364c^3 - 216c^2 + 796c - 2355) > 0 \quad (4.226)$$

egyenlőtlenséget kapjuk. Ez pedig teljesül, hiszen az egyenlőtlenség bal oldala a  $4 < c$  esetben pozitív függvények összege.

Ezzel igazoltuk a tételt. □

A fizikai intuíció alapján várható, hogy a valós ill. komplex sűrűségi mátrixokra megfogalmazott Petz-sejtések ekvivalensek egymással. Ennek igazolása valószínűleg a sejtés bizonyításával összemérhető nehézségű feladat.

### 4.3. Az $\mathcal{M}_2^+$ tér skalárgömbülete lokális minimummal

Az előző rész elején említettük, hogy a monoton metrikával Riemann-sokasággá tett állapotter adott állapotbeli skalárgömbülete szoros összefüggésben áll az adott állapot (környező állapotaitól való) megkülönböztetlenségével. A fizikai intuíció alapján azokat a monoton metrikákat tekintjük fizikailag relevánsaknak vagy elfogadhatónak, melyekből származó skalárgömbület a majorizációra nézve monoton. Ez a gondolatmenet persze megengedi, hogy a majorizációra nézve nem monoton skalárgömbületet adó monoton Riemann-metrikák fontos szerepet játszanak az állapotter statisztikai, információelméleti vagy valamilyen más oldalú megközelítésében. Bizonyos operátormonoton függvényeket a fent említett szempontok alapján csoportosíthatunk.

**4.3. Definíció.** Adott  $n \in \mathbb{N}$  esetén az  $f \in \mathcal{F}_{[0,\infty]}^{(S,n)}$  függvényt és az általa indukált  $K^{(n),f}$  monoton metrikát *n-elfogadhatónak* nevezzük, ha az  $(\mathcal{M}_n^+, K^{(n),f})$  Riemann-sokaság skalárgömbülete a majorizációra nézve monoton. Továbbá az  $f$  függvényről, illetve az általa indukált metrikáról azt mondjuk, hogy *elfogadható*, ha minden  $n \in \mathbb{N}$  esetén *n-elfogadható*. Azon  $f$  függvények halmazát, melyek minden  $k = 1, \dots, n$  esetén *k-elfogadhatók* jelölje  $E_n$ , az elfogadható függvények halmazát pedig  $E$ .

Ezek alapján Petz sejtése úgy fogalmazható meg, hogy az  $f_{\text{KM}}$  függvény elfogadható.

Az eddig bevezetett fontosabb metrikák (melyek a (3.55–3.64) függvényekkel jellemezhetők) közül többről bizonyították, hogy 2-elfogadható. A jelen részben megadunk egy formulát, mellyel az  $\mathcal{M}_2^+$  tér skalárgömbülete egyszerűen származtatható. Numerikus szimulációk szerint az említett fontosabb monoton metrikák mind 2-elfogadhatók. (A P1 és P2 metrikacsalád kivételével ez az előző részben is használt „deriválásos módszerrel” igazolható.)

Felmerül a kérdés, hogy vajon az  $\mathcal{M}_2^+$  téren minden metrika 2-elfogadható lesz-e. Az alábbiakban, Andai [5] cikke alapján bemutatjuk, hogy nem; és példákat adunk nem 2-elfogadható metrikákra.

A (4.80) képlet segítségével meghatározhatjuk az  $(\mathcal{M}_2^+, K^{(n),f})$  tér gömbületét valamely  $D \in \mathcal{M}_2^+$  állapotban. Ebben az esetben, a sokaság „egyszerűsége”

miatt, sokkal kezelhetőbb formulát is nyerhetünk a skalárgörbületre. Ezen formula meghatározásához az egyik legfontosabb segédfüggvény az  $f \in \mathcal{F}_{[0,\infty]}^{(S,n)}$  függvényhez rendelt  $c_f$  Cencov–Morozova-féle függvény lesz.

Jelölje  $\partial_1 n(x, y)$  az  $n$  függvény első változó szerinti parciális deriváltját, és definiáljuk az alábbi négy függvényt [25] alapján először pozitív, páronként különböző  $x, y$  és  $z$  értékekre.

$$h_1(x, y, z) = \frac{c(x, y) - zc(x, z)c(y, z)}{(x - z)(y - z)c(x, z)c(y, z)} \quad (4.227)$$

$$h_2(x, y, z) = \frac{(c(x, z) - c(y, z))^2}{(x - y)^2 c(x, y)c(x, z)c(y, z)}$$

$$h_3(x, y, z) = \frac{z(\partial_1(\log c)(z, x) - \partial_1(\log c)(z, y))}{x - y}$$

$$h_4(x, y, z) = z\partial_1(\log c)(z, x)\partial_1(\log c)(z, y)$$

A  $h_i$  ( $i = 1, 2, 3, 4$ ) függvények értelmezési tartományát ki lehet terjeszteni minden pozitív számhármásra határértékképpéssel. Például a  $h_i(x, x, z)$  értéket a

$$\lim_{y \rightarrow x} h_i(x, y, z) \quad (4.228)$$

kifejezéssel értelmezzük. A  $h_i$  függvények következő lineáris kombinációját fogjuk sokszor használni.

$$h(x, y, z) = h_1(x, y, z) - \frac{1}{2}h_2(x, y, z) + 2h_3(x, y, z) - h_4(x, y, z) \quad (4.229)$$

**4.9. Tétel.** ([25]) Jelölje  $\sigma(D)$  a  $D \in \mathcal{M}_n^+$  állapot sajátértékeinek a halmazát. Ekkor az  $(\mathcal{M}_n^+, K^{(n),f})$  Riemann-geometria skalárgörbületét a  $D$  pontban a

$$\text{Scal}(D) = \sum_{x,y,z \in \sigma(D)} h(x, y, z) - \sum_{x \in \sigma(D)} h(x, x, x) + \frac{1}{4}(n^2 - 1)(n^2 - 2) \quad (4.230)$$

kifejezés adja meg.

**4.10. Tétel.** Az  $(\mathcal{M}_2^+, K^{(2),f})$  tér skalárgörbületre a  $\lambda_1, \lambda_2$  sajátértékekkel rendelkező  $D \in \mathcal{M}_2^+$  állapotban

$$\begin{aligned} \text{Scal}(D) = & h(\lambda_1, \lambda_1, \lambda_2) + h(\lambda_1, \lambda_2, \lambda_1) + h(\lambda_2, \lambda_1, \lambda_1) + h(\lambda_2, \lambda_2, \lambda_1) + \\ & + h(\lambda_2, \lambda_1, \lambda_2) + h(\lambda_1, \lambda_2, \lambda_2) + \frac{3}{2}. \end{aligned} \quad (4.231)$$

**4.11. Tétel.** Legyen a  $D \in \mathcal{M}_2^+$  állapotnak az egyik sajátértéke  $\lambda_1$ . Ekkor  $f \in \mathcal{F}_{[0,\infty]}^{(S,n)}$  esetén az  $(\mathcal{M}_2^+, K^{(2),f})$  tér skalárgörbülete az  $a = 2\lambda_1 - 1$  paraméterrel kifejezve

$$\begin{aligned} Scal(a) = & \frac{14(a-1) \left[ f' \left( \frac{1-a}{1+a} \right) \right]^2}{(1+a)^3 \left[ f \left( \frac{1-a}{1+a} \right) \right]^2} + \frac{2(a^2 + 7a - 6) f' \left( \frac{1-a}{1+a} \right)}{(1+a)^2 a f \left( \frac{1-a}{1+a} \right)} + \frac{8(1-a) f'' \left( \frac{1-a}{1+a} \right)}{(1+a)^3 f \left( \frac{1-a}{1+a} \right)} \\ & + \frac{2(1+a) f \left( \frac{1-a}{1+a} \right)}{a^2} + \frac{3a^3 + 5a^2 + 8a - 4}{2(1+a)a^2}. \end{aligned} \quad (4.232)$$

*Bizonyítás.* A számolás folyamán a minden  $f \in \mathcal{F}_{[0,\infty]}^{(S,n)}$  függvényre érvényes  $f'(1) = \frac{1}{2}$  és  $2f^{(3)}(1) + 3f^{(2)}(1) = 0$  azonosságot fogjuk használni, melyek egyszerűen származtathatók az  $f(x) = xf(1/x)$  és  $f(1) = 1$  egyenlőségből.

A számolások során kihasználjuk a függvények  $h_i(y, x, x) = h_i(x, y, x)$  ( $i = 1, 2, 3, 4$ ) szimmetriatulajdonságát, valamint a Cencov–Morozova-függvényre vonatkozó alábbi azonosságokat.

$$c(x, x) = \frac{1}{x} \quad c(x, y) = c(y, x) \quad c(x, y) = tc(tx, ty), \quad \forall t \in \mathbb{R}^+ \quad (4.233)$$

$$\partial_1 c(x, x) = -\frac{1}{2x^2} \quad \partial_1^k \partial_2^l c(x, y) = \partial_1^l \partial_2^k c(y, x) \quad c(x, y) = -x \partial_1 c(x, y) - y \partial_2 c(x, y)$$

A  $h_i(x, x, y)$  és  $h_i(x, y, x)$  értékeket a fenti azonosságok segítségével határozhatjuk meg. Példaként határozzuk meg a  $h_1(x, x, y)$  és  $h_4(x, y, x)$  értékeket.

$$h_1(x, x, y) = \lim_{q \rightarrow x} h_1(x, q, y) = \lim_{q \rightarrow x} \frac{c(x, q) - yc(x, y)c(q, y)}{(x-y)(q-y)c(x, y)c(q, y)} = \frac{c(x, x) - y[c(x, y)]^2}{(x-y)^2 [c(x, y)]^2} \quad (4.234)$$

$$h_4(x, y, x) = \lim_{q \rightarrow x} h_4(x, y, q) = \lim_{q \rightarrow x} q \frac{\partial_1 c(q, x)}{c(q, x)} \frac{\partial_1 c(q, y)}{c(q, y)} = x \frac{\partial_1 c(x, x)}{c(x, x)} \frac{\partial_1 c(x, y)}{c(x, y)}$$

A (4.233) azonosságok felhasználásával kapjuk a  $h_i$  függvények további speciális értékeit.

$$h_1(x, x, y) = \frac{1 - xy [c(x, y)]^2}{x(x-y)^2 [c(x, y)]^2} \quad h_1(x, y, x) = -\frac{1}{2} \frac{c(x, y) + 2x \partial_1 c(x, y)}{(x-y)c(x, y)} \quad (4.235)$$

$$h_2(x, x, y) = x \left( \frac{\partial_1 c(x, y)}{c(x, y)} \right)^2 \quad h_2(x, y, x) = \frac{1}{x} \left( \frac{1 - xc(x, y)}{(x-y)c(x, y)} \right)^2$$

$$\begin{aligned}
h_3(x, x, y) &= -\frac{y^2 c(x, y) [\partial_1 c(y, x)]^2 + 2yc(x, y) \partial_1 c(y, x) + xy \partial_1 c(x, y) \partial_1 c(y, x)}{x [c(x, y)]^2} \\
h_3(x, y, x) &= -\frac{c(x, y) + 2x \partial_1 c(x, y)}{2(x - y)c(x - y)} \\
h_4(x, x, y) &= y \left( \frac{\partial_1 c(y, x)}{c(x, y)} \right)^2 & h_4(x, y, x) &= -\frac{1}{2} \frac{\partial_1 c(x, y)}{c(x, y)}.
\end{aligned}$$

Minden  $i = 1, 2, 3, 4$  értékre vezessük be az alábbi összegfüggvényt.

$$sh_i(x, y) = h_i(x, x, y) + 2h_i(x, y, x) + h_i(y, y, x) + 2h_i(y, x, y) \quad (4.236)$$

A fentiek alapján a

$$sh_1(x, y) = \frac{(x + y)(1 - xy [c(x, y)]^2)}{xy(x - y)^2 [c(x, y)]^2} - \frac{4x \partial_1 c(x, y) + 2c(x, y)}{(x - y)c(x, y)} \quad (4.237)$$

$$sh_2(x, y) = \frac{x [\partial_1 c(x, y)]^2 + y [\partial_1 c(y, x)]^2}{[c(x, y)]^2} + 2 \frac{(x + y) + xy(x + y) [c(x, y)]^2 - 4xyc(x, y)}{xy(x - y)^2 [c(x, y)]^2}$$

$$sh_3(x, y) = \frac{(x + y)(c(x, y) [\partial_{1,2} c(x, y)]^2 - \partial_1 c(x, y) \partial_1 c(y, x))}{[c(x, y)]^2} - \frac{4x \partial_1 c(x, y) + 2c(x, y)}{(x - y)c(x, y)}$$

$$sh_4(x, y) = \frac{x [\partial_1 c(x, y)]^2 + y [\partial_1 c(y, x)]^2}{[c(x, y)]^2} - \frac{\partial_1 c(x, y) + \partial_1 c(y, x)}{c(x, y)}$$

kifejezéseket kapjuk.

Ezeket az összegfüggvényeket kifejezhetjük az operátormonoton  $f$  függvénnyel.

$$sh_1(x, y) = \frac{1}{(x - y)^2} \left( \frac{y(x + y)}{x} [f(x/y)]^2 + y - 3x + \frac{4x(x - y)}{y} \frac{f'(x/y)}{f(x/y)} \right) \quad (4.238)$$

$$\begin{aligned}
sh_2(x, y) &= \frac{x}{y^2} \left( \frac{f'(x/y)}{f(x/y)} \right)^2 + \frac{y^3}{x^4} \left( \frac{f(x/y)f'(y/x)}{[f(y/x)]^2} \right)^2 + \frac{2y(x + y)}{x(x - y)^2} [f(x/y)]^2 - \\
&\quad - \frac{8y}{(x - y)^2} f(x/y) + \frac{2(x + y)}{(x - y)^2}
\end{aligned}$$

$$\begin{aligned}
sh_3(x, y) &= \frac{-2x(x + y)}{y^3} \left( \frac{f'(x/y)}{f(x/y)} \right)^2 - \frac{(x + y)}{x^2} \frac{f'(x/y)f'(y/x)}{[f(x/y)]^2} + \\
&\quad + \frac{2(x^2 + 2xy - y^2)}{y^2(x - y)} \frac{f'(x/y)}{f(x/y)} + \frac{x(x + y)}{y^3} \frac{f''(x/y)}{f(x/y)} - \frac{2}{x - y}
\end{aligned}$$

$$sh_4(x, y) = \frac{x}{y^2} \left( \frac{f'(x/y)}{f(x/y)} \right)^2 + \frac{y^3}{x^4} \left( \frac{f(x/y)f'(y/x)}{[f(y/x)]^2} \right)^2 + \frac{1}{y} \frac{f'(x/y)}{f(x/y)} + \frac{y}{x^2} \frac{f(x/y)f'(y/x)}{[f(y/x)]^2}$$

A skalárgörbületet az összegfüggvények alábbi lineáris kombinációja adja.

$$\text{Scal}(D) = sh_1(x, y) - \frac{1}{2}sh_2(x, y) + 2sh_3(x, y) - sh_4(x, y) + \frac{3}{2} \quad (4.239)$$

Ezeket egybevetve kapjuk a

$$\begin{aligned} \text{Scal}(D) = & 2 \frac{2yf(x/y) - 1}{(x-y)^2} + 6 \frac{2xf'(x/y) - yf(x/y)}{y(x-y)f(x/y)} - \frac{1}{2} \frac{x(8+3y)}{y^3} \left( \frac{f'(x/y)}{f(x/y)} \right)^2 + \frac{3}{2} - \\ & - \frac{3}{2} \frac{y}{x^2} \left( \frac{f'(y/x)}{f(y/x)} \right)^2 + \frac{(3+x)f'(x/y)}{y^2 f(x/y)} + 2 \frac{xf''(x/y)}{y^3 f(x/y)} - \frac{f'(y/x)}{xf(y/x)} - \frac{2f'(x/y)f'(y/x)}{x^2 [f(y/x)]^2} \end{aligned} \quad (4.240)$$

kifejezést a skalárgörbületre.

A  $D$  állapot sajátértékeit ki lehet fejezni az  $a$  paraméterrel

$$\lambda_1 = \frac{1+a}{2} \quad \lambda_2 = \frac{1-a}{2} . \quad (4.241)$$

A sajátértékek ezen formájának a fenti egyenletbe való helyettesítésével kapjuk a bizonyítandó (4.232) kifejezést.  $\square$

Mivel a skalárgörbületre vonatkozó (4.232) kifejezés meglehetősen bonyolult, ezért érdemes példaként megemlíteni egy másik – differenciálgeometriai – bizonyítást, mely nem támaszkodik Dittmann eredményére.

**4.1. Példa.** A 3.1. példában bemutattuk az  $\mathcal{M}_2^+$  tér Stokes-paraméterezését, melynek a segítségével az  $\mathcal{M}_2^+$  tér azonosítható a háromdimenziós nyílt egységgömbbel. Ezen az egységgömbön a gömbi koordinátarendszer bevezetésével az  $\mathcal{M}_2^+$  állapotér

$$\frac{1}{2} \begin{pmatrix} 1 + r \cos \theta & (r \sin \theta \cos \phi) + i(r \sin \theta \sin \phi) \\ (r \sin \theta \cos \phi) - i(r \sin \theta \sin \phi) & 1 - r \cos \theta \end{pmatrix} \quad (4.242)$$

paraméterezését kapjuk, ahol  $(r, \phi, \theta)$  jelöli a szokásos gömbi koordinátákat, de a jelen esetben  $0 \leq r < 1$ . Ezt a paraméterezést az  $\mathcal{M}_2^+$  tér *gömbi paraméterezésének* nevezzük. A továbbiakban legyen a lokális koordináták sorrendje  $(r, \theta, \phi)$ . (Így például a koordinátáktól függő  $t$  függvény esetén  $\partial_2 t$  jelöli a  $t$  függvény  $\theta$  változó szerinti parciális deriváltját.)

A  $g = K^{(2),f}$  metrika ezzel a koordinátázással

$$g_{11} = \frac{1}{1-r^2} \quad (4.243)$$

$$g_{22} = \frac{r^2}{(1+r)f\left(\frac{1-r}{1+r}\right)}$$

$$g_{33} = \frac{r^2 \sin^2 \theta}{(1+r)f\left(\frac{1-r}{1+r}\right)} .$$

A  $g_{ik}$  metrika felírható a  $g_{ik} = \delta_{ik}\alpha_i(r, \theta, \phi)$  alakban. Az

$$\partial_2\alpha_1 = \partial_3\alpha_1 = 0, \quad \partial_2\alpha_2 = \partial_3\alpha_2 = 0, \quad \partial_3\alpha_3 = 0 \quad (4.244)$$

azonosságok felhasználásával egyszerűbben lehet kiszámolni a különböző differenciál-geometriai mennyiségeket.

A (2.31) képlettel értelmezett másodfajú Christoffel-szimbólumot a

$$\Gamma_{ij}^{..m} = \sum_{k=1}^3 \frac{1}{2} g^{km} (\partial_i g_{jk} + \partial_j g_{ik} - \partial_k g_{ij}) \quad (4.245)$$

kifejezés adja, ahol  $g^{ij}$  jelöli a  $g_{ij}$  mátrix inverzét.

A  $\Gamma_{ij}^{..m} = \Gamma_{ji}^{..m}$  szimmetria miatt összesen hét független, nem nulla másodfajú Christoffel-szimbólum lesz ebben az esetben. Ezeket a

$$\Gamma_{1,1}^{..1} = \frac{r}{1-r^2} \quad \Gamma_{2,2}^{..1} = -\frac{r(1-r)}{2(1+r)^2 f(c(r))} \left( r^2 + 3r + 2 + 2r \frac{f'(c(r))}{f(c(r))} \right) \quad (4.246)$$

$$\Gamma_{3,3}^{..1} = \sin^2 \theta \Gamma_{2,2}^{..1} \quad \Gamma_{1,2}^{..2} = -\frac{f(c(r))}{r^2(1-r)} \Gamma_{2,2}^{..1} \quad \Gamma_{3,3}^{..2} = -\sin \theta \cos \theta$$

$$\Gamma_{1,3}^{..3} = \Gamma_{1,2}^{..2} \quad \Gamma_{2,3}^{..3} = \frac{\cos \theta}{\sin \theta}$$

kifejezés adja meg, ahol  $c(r) = \frac{1-r}{1+r}$ .

A (2.39) képletnek megfelelően a Riemann-féle görbületi tenzort a

$$R_{ijkl} = \sum_{n=1}^3 g_{ln} \left( \partial_i \Gamma_{jk}^{..n} - \partial_j \Gamma_{ik}^{..n} + \sum_{m=1}^3 (\Gamma_{jk}^{..m} \Gamma_{im}^{..n} - \Gamma_{ik}^{..m} \Gamma_{jm}^{..n}) \right) \quad (4.247)$$

képlettel határozhatjuk meg. Ezen tenzornak  $R_{ijkl} = -R_{jikl}$ ,  $R_{ijkl} = -R_{ijlk}$  és  $R_{ijkl} = R_{klij}$  szimmetriatulajdonságai miatt, összesen három független, nem nulla eleme lesz.

$$R_{1212} = -\frac{r}{(1+r)^4(1-r^2)f(c(r))} \left( 2r(1-r) \frac{f''(c(r))}{f(c(r))} - 3r(1-r) \left( \frac{f'(c(r))}{f(c(r))} \right)^2 \right) \quad (4.248)$$

$$+(1+r)(3r-2) \frac{f'(c(r))}{f(c(r))} + \frac{(r^2+r+4)(1+r)^2}{4}$$

$$R_{1313} = \sin^2 \theta R_{1212}$$

$$R_{2323} = \frac{r^2(1-r)\sin^2\theta}{(1+r)^4 [f(c(r))]^2} \left( r(r+2) \frac{f'(c(r))}{f(c(r))} + \frac{r^2}{1+r} \left( \frac{f'(c(r))}{f(c(r))} \right)^2 - \frac{(1+r)^3}{1-r} f(c(r)) \right. \\ \left. + \frac{(1+r)(2+r)^2}{4} \right)$$

A (2.42) egyenlet alapján a Ricci-tenzort a

$$\text{Ric}_{ij} = \sum_{k,l=1}^3 g^{kl} R_{lijk}. \quad (4.249)$$

képlettel határozhatjuk meg. A  $\text{Ric}_{ij} = \text{Ric}_{ji}$  tulajdonság miatt ennek a tenzornak is csak három független, nem nulla eleme lesz.

$$\text{Ric}_{1,1} = \frac{1}{(1+r)^4} \left( 4 \frac{f''(c(r))}{f(c(r))} - 6 \left( \frac{f'(c(r))}{f(c(r))} \right)^2 + \frac{2(1+r)(3r-2)}{(1-r)} \frac{f'(c(r))}{f(c(r))} \right. \\ \left. + \frac{(r^2+r+4)(1+r)^2}{2r(1-r)} \right) \quad (4.250)$$

$$\text{Ric}_{2,2} = \frac{r^2(1-r)}{(1+r)^4 f(c(r))} \left( 2 \frac{f''(c(r))}{f(c(r))} - 4 \left( \frac{f'(c(r))}{f(c(r))} \right)^2 + \frac{(1+r)(r^2+4r-4)}{r(1-r)} \frac{f'(c(r))}{f(c(r))} \right. \\ \left. + \frac{(1+r)^4}{r^2(1-r)} f(c(r)) + \frac{(r^3+2r^2+2r-2)(1+r)^2}{2r^2(1-r)} \right)$$

$$\text{Ric}_{3,3} = \sin^2\theta \text{ Ric}_{2,2}.$$

A (2.49) képlet alapján az  $(\mathcal{M}_2^+, K^{(2),f})$  Riemann-geometria  $D \in \mathcal{M}_2^+$  pontjában a skalárgörbület

$$\text{Scal}(D) = \sum_{i,j=1}^3 g^{ij} \text{Ric}_{ji}. \quad (4.251)$$

A fenti összegzés elvégzésével visszakapjuk a (4.232) képletet.

□

**4.2. Példa.** A 4.10. tétel alapján meghatározhatjuk a főbb operátormonoton függvények által generált  $(\tilde{\mathcal{M}}_n^+, \tilde{K}^{(n),f})$  Riemann-metrikák skalárgörbületét.



Az  $f_{SM}$  függvény által generált legkisebb monoton metrika esetén a skalárgörbület állandó

$$\text{Scal}_{SM}(a) = \frac{9}{2}. \quad (4.252)$$

Az  $f_{LA}$  függvény által generált legnagyobb monoton metrika esetén a skalárgörbület

$$\text{Scal}_{SM}(a) = -\frac{7a^2 - 27}{2(1+a)(a-1)}. \quad (4.253)$$

Az  $f_{KM}$  függvény által generált Kubo–Mori-metrika esetén

$$\text{Scal}_{KM}(a) = \frac{1}{2a^2} + 2\frac{1+a^2}{a(1-a)(1+a)\log\left(\frac{1-a}{1+a}\right)} + 2\frac{1}{(1+a)(1-a)\left(\log\left(\frac{1-a}{1+a}\right)\right)^2}. \quad (4.254)$$

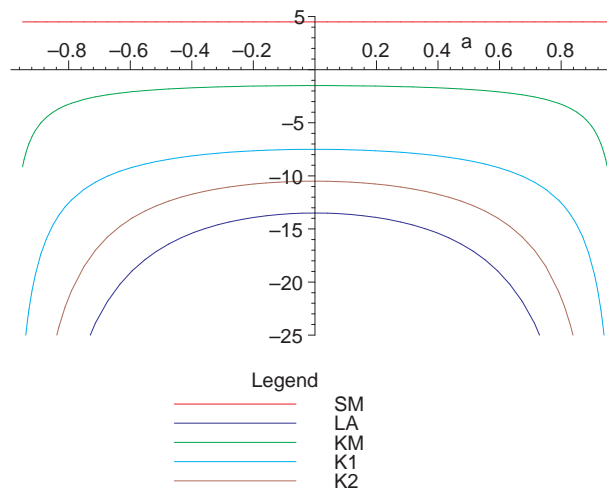
Az  $f_{K1}$  függvény által generált metrika esetén

$$\text{Scal}_{K1}(a) = -\frac{3}{2} + 8\frac{a}{(1-a)(1+a)\log\left(\frac{1-a}{1+a}\right)} - 8\frac{a^2}{(1+a)(1-a)\left(\log\left(\frac{1-a}{1+a}\right)\right)^2}. \quad (4.255)$$

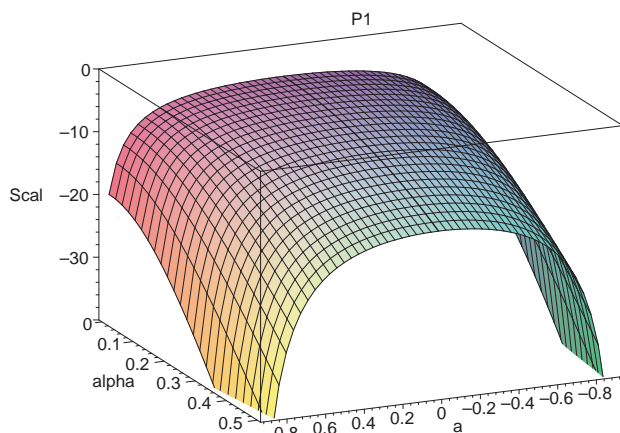
Az  $f_{K2}$  függvény által generált metrika esetén

$$\text{Scal}_{K2}(a) = -\frac{3a^4 - 11a^2 + 1}{2a^2(1-a)(1+a)} - 2\frac{2\sqrt{1-a^2} - \frac{2a^2+3}{1-a^2}}{a\log\left(\frac{1-a}{1+a}\right)} + 2\frac{1}{(1+a)(1-a)\left(\log\left(\frac{1-a}{1+a}\right)\right)^2}. \quad (4.256)$$

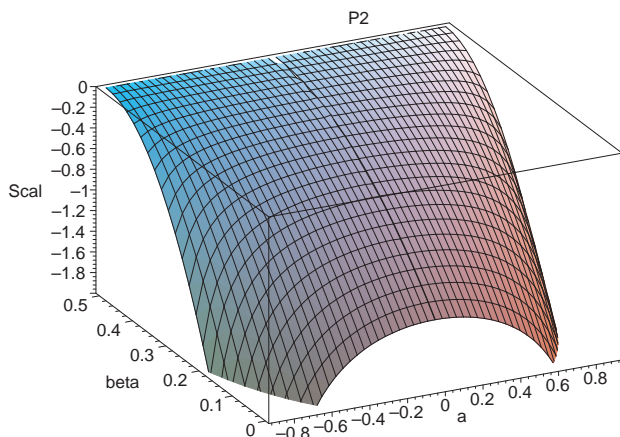
Ezekben az esetekben a skalárgörbületnek lokális maximuma van a legkevesebb állapotban, mint ahogy az az alábbi ábrán látható.



Az  $f_{P1}$  és  $f_{P2}$  függvény esetén meglehetősen bonyolult kifejezés adódik a skalárgörbületre. Ezért csak az alábbi ábrákon mutatjuk be, hogy a  $0 \leq \alpha \leq \frac{1}{2}$  paramétertartományban hogyan változik a skalárgörbület,



illetve  $0 \leq \beta \leq \frac{1}{2}$  esetén.



Figyeljük meg a második ábránál, hogy  $\beta = \frac{1}{2}$  esetén a skalárgörbület azonosan nulla.

□

A  $D \in \mathcal{M}_2^+$  állapot akkor a legkevertebb, ha a két sajátértéke megegyezik, ebben az esetben  $a = 0$ . Vagyis a skalárgörbületnek akkor van lokális minimuma vagy maximuma a legkevertebb állapotban, ha a (4.232) képlettel értelmezett Scal függvénynek lokális minimuma vagy maximuma van az origóban.

A Scal függvény Taylor-sorfejtésének a segítségével vizsgáljuk a skalárgörbület lokális tulajdonságait a legkevertebb állapot közelében.

**4.12. Tétel.** A (4.232) képlettel értelmezett Scal függvény Taylor-sorfejtése az origóban a

$$r(a) = (6 + 36f''(1)) + a^2 \left( \frac{100}{3}f^{(4)}(1) - 140f''(1) - 120f''(1)^2 \right) + a^4 \left( 352f''(1)^3 + \right. \\ \left. + 616f''(1)^2 + 1092f''(1) - \frac{1288}{3}f^{(4)}(1) + \frac{392}{45}f^{(6)}(1) - 160f''(1)f^{(4)}(1) \right) + O(a^6) \quad (4.257)$$

közelítő formulát adja.

*Bizonyítás.* Az (4.232) egyenlet alapján várható lenne, hogy  $\frac{1}{a}$  és  $\frac{1}{a^2}$  típusú divergenciák is megjelennek a sorfejtésben, azonban az operátormonoton  $f$  függvény deriváltjainak a viselkedése ezt nem engedi.

A skalárgörbületre nyilván  $\text{Scal}(a) = \text{Scal}(-a)$  teljesül (nem a (4.232) képlet miatt) szimmetria okok miatt, ezért a sorfejtésben csak páros rendű tagok jelennek meg. Vagyis az  $a^{2n+1}$  alakú tagok együtthatója nulla minden  $a \in \mathbb{N}$  esetén.

A sorfejtést csak másodrendig bizonyítjuk be, teljesen hasonló módon, csak jóval több számolással bizonyítható a negyedrendű tag is. A számolás során a minden  $f \in \mathcal{F}_{[0,\infty]}^{(S,n)}$  függvényre érvényes

$$f'(1) = \frac{1}{2}, \quad f^{(3)}(1) = -\frac{3}{2}f^{(2)}(1), \quad f^{(5)}(1) = -\frac{15f^{(4)}(1) + 60f^{(3)}(1) + 60f^{(2)}(1)}{2} \quad (4.258)$$

azonosságokat fogjuk használni, melyek egyszerűen származtathatók a  $f(x) = xf(1/x)$  és  $f(1) = 1$  egyenletből.

A  $\text{Scal}$  függvényt, a (4.232) képletnek megfelelően öt függvény összegeként tekintjük. Minden egyes összeadandó sorfejtése az  $a = 0$  pont körül elemi számolással meghatározható, azonban a részeredmények meglehetősen bonyolult képleteket tartalmaznak. Egyszerűsítések után az alábbiakat kapjuk az összeadandók sorfejtésére.

$$1. \text{ tag: } -\frac{7}{2} + 7(4f''(1) + 1) \cdot a - 7(8f''(1)^2 + 4f''(1) + 1) \cdot a^2 \quad (4.259)$$

$$2. \text{ tag: } -6 \cdot \frac{1}{a} + (24f''(1) + 13) - (28f''(1) + 12) \cdot a + \\ + (16f^{(4)}(1) - 48f''(1)^2 - 52f''(1) + 12) \cdot a^2$$

$$3. \text{ tag: } 8f''(1) + (16f^{(4)}(1) - 16f''(1)^2 - 56f''(1)) \cdot a^2$$

$$4. \text{ tag: } 2 \cdot \frac{1}{a^2} + 4f''(1) + \left(\frac{4}{3}f^{(4)}(1) - 4f''(1)\right) \cdot a^2$$

$$5. \text{ tag: } -2 \cdot \frac{1}{a^2} + 6 \cdot \frac{1}{a} - 5 + 5 \cdot a - 5 \cdot a^2$$

Ezen öt tag összege adja a tételben szereplő sorfejtést másodrendig.  $\square$

Ahhoz, hogy találjunk olyan  $f \in \mathcal{F}_{[0,\infty]}^{(S,n)}$  függvényt, melyre a skalárgörbületnek lokális minimuma van a legkevertebb állapotban, az operátormonoton függvényekre vonatkozó 3.11. reprezentációs tételt használjuk.

A (3.43) és (4.257) egyenletekből adódik, hogy az  $(\mathcal{M}_n^+, K^{(2),f})$  sokaság skalárgörbületének akkor van lokális minimuma a legkevertebb állapotban, ha teljesül az

$$12 \left( \int_0^1 t(1-t) d\mu(t) \right)^2 - \int_0^1 t(t-1)(20t^2 - 40t + 13) d\mu(t) < 0 \quad (4.260)$$

egyenlőtlenség, ahol a  $\mu \in \mathcal{G}_{[0,1]}^{(S,n)}$  mérték határozza meg az operátormonoton  $f$  függvényt a (3.43) képlettel. Továbbá a skalárgörbület értéke a legkevertebb állapotban

$$\text{Scal}(0) = 6 + 72 \int_0^1 (t^2 - t) d\mu(t) . \quad (4.261)$$

A  $\text{Scal}(0)$  érték a  $\mu = (1/2)\delta_0 + (1/2)\delta_1$  mérték esetén a legnagyobb, ekkor  $\text{Scal}(0) = 6$  és ez a mérték az  $f_{\text{SM}}$  függvényt generálja. A  $\text{Scal}(0)$  a legkisebb értékét a  $\mu = \delta_{1/2}$  mérték esetén veszi fel, ekkor  $\text{Scal}(0) = -12$  és ez a mérték az  $f_{\text{LA}}$  függvényt generálja.

Minden  $\mu \in \mathcal{G}_{[0,1]}^{(S,n)}$  mértéket át lehet transzformálni egy  $\mu' \in \mathcal{G}_{[0,1]}$  valószínűségi mértékké úgy, hogy minden  $\varphi : [0, 1] \rightarrow \mathbb{R}$  mérhető függvényre

$$\int_0^1 \varphi(x) d\mu'(x) = 2 \int_0^{\frac{1}{2}} \varphi(4t(1-t)) d\mu(t) \quad (4.262)$$

teljesül, ugyanis a  $t \mapsto 4t(1-t)$  függvény kölcsönösen egyértelmű megfeleltetést létesít a  $[0, \frac{1}{2}]$  és a  $[0, 1]$  intervallum között.

Ha  $\lambda$  jelöli a Lebesgue-mértéket és

$$\mu(t)|_{[0, \frac{1}{2}]} = \rho(t) d\lambda(t) + \sum a_i \delta_{p_i} , \quad (4.263)$$

akkor

$$\mu'(x) = \frac{1}{2} \rho \left( \frac{1 - \sqrt{1-x}}{2} \right) \frac{1}{\sqrt{1-x}} d\lambda(x) + \sum 2a_i \delta_{4p_i(1-p_i)} . \quad (4.264)$$

Így kölcsönösen egyértelmű megfeleltetés létesíthető a  $[0, 1]$  intervallumon értelmezett valószínűségi mértékek és  $\mathcal{G}_{[0,1]}^{(S,n)}$  elemei között.

Adott  $\mu \in \mathcal{G}_{[0,1]}^{(S,n)}$  mérték esetén jelölje  $\mu'$  a neki megfelelő  $[0, 1]$  intervallumon értelmezett valószínűségi mértéket,  $m_\mu$  ennek várható értékét,  $\sigma_\mu^2$  a varianciáját és  $E_{n,\mu}$  pedig az  $n$ -edik momentumát. A (3.43) egyenlet és az imént bevezetett jelölések alapján az

$$f''(1) = -\frac{m_\mu}{2} \quad f^{(4)}(1) = -3m_\mu + \frac{3}{2}E_{2,\mu} \quad f^{(6)}(1) = -90m_\mu - \frac{45}{4}E_{3,\mu} + 90E_{2,\mu} \quad (4.265)$$

egyenlőségeket kapjuk.

Ezt a (4.257) sorfejtésbe helyettesítve kapjuk a következő tételt.

**4.13. Tétel.** Ha a  $\mu \in \mathcal{G}_{[0,1]}^{(S,n)}$  mértékre

$$m_\mu(3 - 2m_\mu) < 5\sigma_\mu^2 \quad (4.266)$$

teljesül, vagy ha

$$m_\mu(3 - 2m_\mu) = 5\sigma_\mu^2 \quad \text{és} \quad -44m_\mu^3 + 70m_\mu^2 + 114m_\mu < 98E_{3,\mu}, \quad (4.267)$$

teljesül, akkor a  $\mu$  mérték által a (3.43) képlettel indukált  $f$  függvényre, az  $(\mathcal{M}_2^+, K^{(2),f})$  Riemann-sokaság skalárgörbületének lokális minimuma lesz a legkevertebb állapotban.

A továbbiakban példákat adunk olyan  $f \in \mathcal{F}_{[0,\infty]}^{(S,n)}$  függvényekre, melyekre az  $(\mathcal{M}_2^+, K^{(2),f})$  tér skalárgörbületének lokális minimuma van a legkevertebb állapotban.

**4.14. Tétel.** Legyen  $p \in \mathbb{R}$  olyan paraméter, melyre

$$\frac{7 - \sqrt{7}}{14} < p \leq \frac{1}{2} \quad (4.268)$$

teljesül, valamint  $q \in \mathbb{R}$  legyen olyan paraméter, melyre  $0 \leq q < \frac{1}{2} - h(p)$ , ahol

$$h(p) = \frac{\sqrt{14p^2 - 14p + 4 + \sqrt{-640p^4 + 1280p^3 - 880p^2 + 240p + 9}}}{2\sqrt{7}}. \quad (4.269)$$

Ekkor az

$$f(x) = \frac{x}{4} \left( \frac{1}{px + 1 - p} + \frac{1}{(1-p)x + p} + \frac{1}{qx + 1 - q} + \frac{1}{(1-q)x + q} \right) \quad (4.270)$$

függvény  $\mathcal{F}_{[0,\infty]}^{(S,n)}$ -beli, és az  $(\mathcal{M}_2^+, K^{(2),f})$  tér skalárgörbületének lokális minimuma van a legkevertebb állapotban.

*Bizonyítás.*

A 4.13. tétel (4.260) feltételének eleget tevő  $\mu \in \mathcal{G}_{[0,1]}^{(S,n)}$  mértéket először keressük

$$\mu_p = \frac{1}{2}\delta_p + \frac{1}{2}\delta_{1-p} \quad (4.271)$$

alakban. Vezessük be a

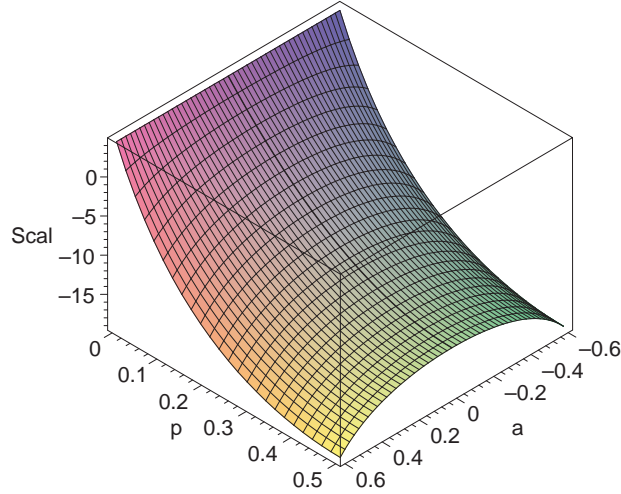
$$t_\mu = 12 \left( \int_0^1 t(1-t) d\mu(t) \right)^2 - \int_0^1 t(t-1)(20t^2 - 40t + 13) d\mu(t) \quad (4.272)$$

jelölést és legyen  $t(p) = t_{\mu_p}$ . Ekkor az

$$t(p) = p(1-p)(8p^2 - 8p + 3) \quad (4.273)$$

egyenlőséget kapjuk. Ezért minden  $p \in [0, 1/2]$  értékre  $t(p) > 0$ . Ez azt jelenti, hogy minden  $\mu_p$  mérték esetén a skalárgörbületnek lokális maximuma lesz az origóban.

Az alábbi ábra szemlélteti, hogy a  $0 \leq p \leq \frac{1}{2}$  paraméter esetén, hogy alakul a  $\text{Scal}(a)$  függvény.



Most próbáljuk  $\mu \in \mathcal{G}_{[0,1]}^{(S,n)}$  mértéket négy Dirac-mértékből összerakni. A  $p \in [0, 1/2]$  és  $q \in [p, 1/2]$  paraméterre legyen

$$\mu_{p,q} = \frac{1}{4}\delta_p + \frac{1}{4}\delta_q + \frac{1}{4}\delta_{1-p} + \frac{1}{4}\delta_{1-q} . \quad (4.274)$$

A  $t(p, q) = t_{\mu_{p,q}}$  jelöléssel az alábbi adódik.

$$t(p, q) = -7(p^4 + q^4) + 14(p^3 + q^3) - 6pq(p + q - pq - 1) - \frac{17}{2}(p^2 + q^2) + \frac{3}{2}(p + q) \quad (4.275)$$

A

$$p = \frac{v + \sqrt{v^2 - 4u}}{2} \quad q = \frac{v - \sqrt{v^2 - 4u}}{2} \quad (4.276)$$

helyettesítést elvégezve a (4.275) egyenletben a

$$t(u, v) = -8u^2 + (28v^2 - 48v + 23)u - \left(7v^4 - 14v^3 + \frac{17}{2}v^2 - \frac{3}{2}v\right) \quad (4.277)$$

kifejezést kapjuk a  $t$  függvényre.

Adott  $v$  érték esetén két megoldása a  $t(u, v) = 0$  egyenletnek. A  $p$  és  $q$  változóra tett feltevések és az  $u = pq$  azonosság miatt a  $0 < u < \frac{1}{4}$  feltételt kapjuk  $u$ -ra. Adott  $v$

érték esetén a  $t(u, v) = 0$  egyenlet egyetlen olyan  $u$  megoldása, mely eleget tesz a fenti feltételnek az alábbi.

$$u(v) = \frac{7}{4}v^2 - 3v + \frac{23}{16} - \frac{1}{16}\sqrt{560v^4 - 2240v^3 + 3320v^2 - 2160v + 529} \quad (4.278)$$

Adott  $p$  paraméter esetén  $q$ -t meg lehet határozni az  $u(p+q) = pq$  egyenletből. Így négy megoldás fog adódni  $q$ -ra, azonban csak egy teljesíti a  $q$ -ra megkövetelt feltételeket, a következő

$$w(p) = \frac{1}{2} - \frac{1}{14}\sqrt{84p^2 - 84p + 28 + 7\sqrt{-640p^4 + 1280p^3 - 880p^2 + 240p + 9}}. \quad (4.279)$$

Ez az egyenlet akkor ad pozitív  $q$  értéket, ha

$$\frac{7 - \sqrt{7}}{14} < p \leq \frac{1}{2} \quad (4.280)$$

teljesül. Könnyen ellenőrizhető, hogy ha  $0 < q < w(p)$ , akkor a  $t(p, q)$  függvény negatív. Ezután a (3.43) egyenlettel definiálhatjuk a  $\mu_{p,q} \in \mathcal{G}_{[0,1]}^{(S,n)}$  mérték által meghatározott  $f \in \mathcal{F}_{[0,\infty]}^{(S,n)}$  függvényt.  $\square$

Ha a  $p$  értéket tetszőlegesen választjuk a  $\left] \frac{7-\sqrt{7}}{14}, \frac{1}{2} \right]$  intervallumból és  $q = 0$ , akkor az előző tétel alapján az

$$f_p(x) = \frac{x}{4} \left( \frac{1}{(1-p)x + p} + \frac{1}{px + 1 - p} + \frac{1}{x} + 1 \right) \quad (4.281)$$

operátormonoton függvény által indukált monoton metrika esetén a skalárgörbületnek lokális minimuma lesz a legkevertebb állapotban. Ekkor a skalárgörbület (4.232) formájának a sorfejtése a legkevertebb állapotban

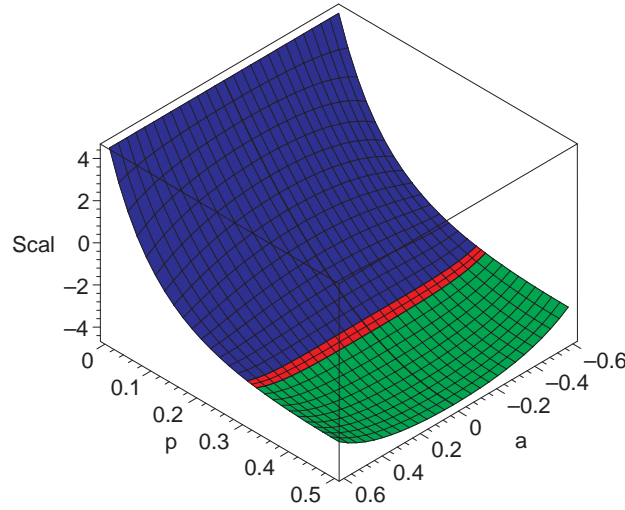
$$\text{Scal}(a) = \left( \frac{9}{2} - 36p(1-p) \right) - 20p(1-p)(14p^2 - 14p + 3) \cdot a^2 + O(a^4). \quad (4.282)$$

Igazolható, hogy ekkor a skalárgörbületnek ekkor nem csak lokális minimuma van az  $a = 0$  helyen, hanem ez egyben globális minimum is. A skalárgörbület-függvény értékkészletének a szuprénuma ekkor

$$\text{Scal}(1) = \frac{7}{2} + \frac{1}{p(1-p)}, \quad (4.283)$$

ahol a  $\text{Scal}(1)$  kifejezés csak formálisan szerepel, hiszen az állapottér a nyílt egység-gömb.

Az alábbi ábra mutatja, hogy változik az  $f_p$  függvény által indukált skalárgörbület a  $0 \leq p \leq \frac{1}{2}$  paramétertartományban. A zölddel jelölt paramétertartományban lokális minimuma, a kézzel jelölt tartományban lokális maximuma van a skalárgörbületnek a legkevésbé állapotban.



Példaként említjük az alábbi operátormonoton függvényeket, melyek által indukált skalárgörbületnek lokális-, de nem globális minimuma van a legkevésbé állapotban.

$$f(x) = \frac{x}{4} \left( \frac{4}{x+1} + \frac{50}{x+49} + \frac{50}{49x+1} \right), \quad f(x) = \frac{250x}{999x+1} + \frac{250x}{x+999} + \frac{x}{x+1} \quad (4.284)$$

A 4.14. tétel bizonyítása közben láttuk, hogy létezik egy folytonos út az  $f_{SM}$  és az  $f_{LA}$  függvények között, a  $\mu_p = \frac{1}{2}\delta_p + \frac{1}{2}\delta_{1-p}$  ( $0 \leq p \leq \frac{1}{2}$ )  $\mathcal{G}_{[0,1]}^{(S,n)}$ -beli mértékek és a (3.43) reprezentációs tétel segítségével. A  $\mu_p$  mérték által generált operátormonoton függvények családját jelölje  $f_{D1}$ , azaz

$$f_{D1}^{(p)}(x) = \frac{x}{2} \left( \frac{1}{px+1-p} + \frac{1}{(1-p)x+p} \right). \quad (4.285)$$

Ekkor

$$f_{SM} = f_{D1}^{(p=0)} \underset{[0,1]}{\geq} f_{D1}^{(0 \leq p \leq 1/2)} \underset{[0,1]}{\geq} f_{D1}^{(p=1/2)} = f_{LA} \quad (4.286)$$

teljesül. Az  $f_{D1}$  függvénycsalád duálisát jelölje  $f_{DI}$ , azaz

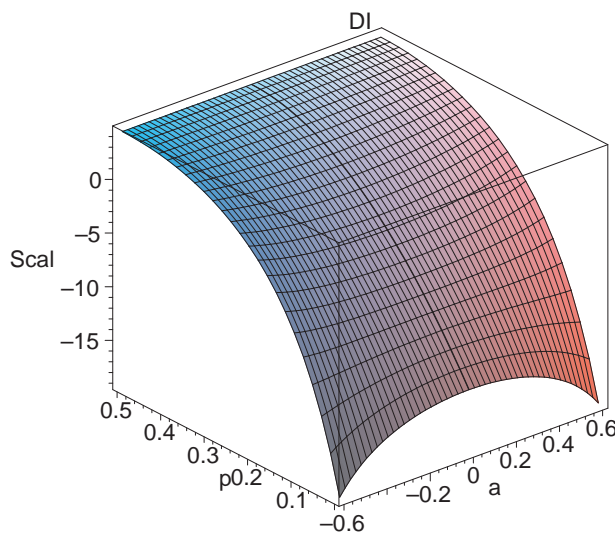
$$f_{DI}^{(p)}(x) = \frac{2}{x+1} (px+1-p)((1-p)x+p) \quad 0 \leq p \leq \frac{1}{2}. \quad (4.287)$$



Ekkor egy újabb utat adhatunk meg a legnagyobb illetve a legkisebb függvény között.

$$f_{\text{SM}} = f_{\text{DI}}^{(p=1/2)} \underset{[0,1]}{\geq} f_{\text{DI}}^{(0 \leq p \leq 1/2)} \underset{[0,1]}{\geq} f_{\text{DI}}^{(p=0)} = f_{\text{LA}} \quad (4.288)$$

Ezen függvénycsaládhoz tartozó skalárgörbület az alábbi ábra alapján szintén 2-elfogadhatónak tűnik.

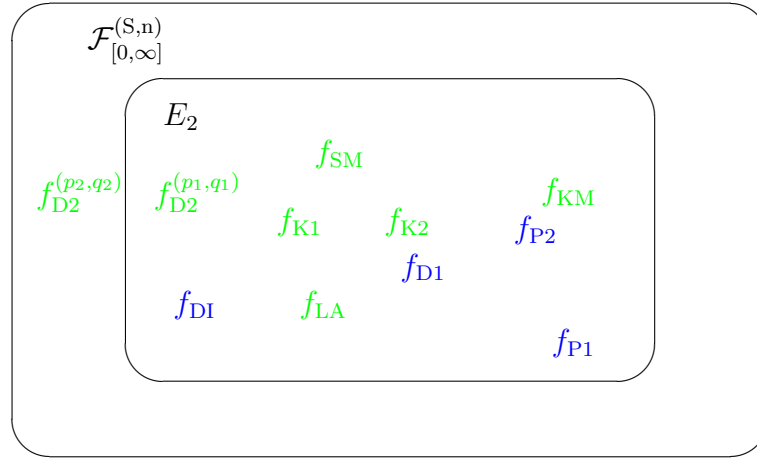


Jelölje  $f_{\text{D2}}$  a  $0 \leq p \leq \frac{1}{2}$  és  $0 \leq q \leq p$  paraméterhez tartozó

$$f_{\text{D2}}^{(p,q)}(x) = \frac{x}{4} \left( \frac{1}{px + 1 - p} + \frac{1}{(1-p)x + p} + \frac{1}{qx + 1 - q} + \frac{1}{(1-q)x + q} \right) \quad (4.289)$$

operátormonoton függvényt.

Az alábbi ábrán foglalhatjuk össze a jelen részben elért eredményeket. (Vagyis nem 2-elfogadható metrikák létezését, és a főbb metrikák 2-elfogadhatóságát.) A zöld színnel jelzett metrikákról matematikailag bizonyítható, hogy abba a halmazba tartoznak, ahol az ábrán szerepelnek, a kék színnel jelzettekről csak numerikus szimulációk állnak rendelkezésre. Metrikák családjá esetén, ha nincs kiírva a paraméter (például  $f_{\text{DI}}$ ), akkor úgy értjük, hogy minden paraméter esetén az adott halmazban van, ha ki van írva a paraméter (például  $f_{\text{D2}}^{(p_1, q_1)}$ ), akkor úgy értjük, hogy az adott paraméterhez tartozó metrika van az adott halmazban.



#### 4.4. Az állapotér skalárgömbületének numerikus vizsgálata

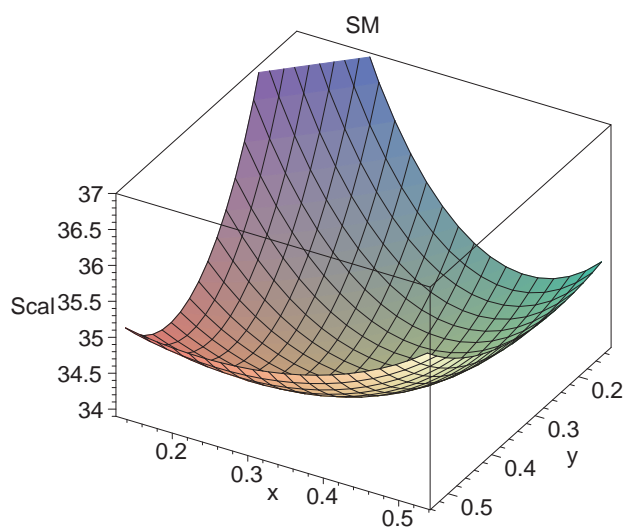
Az előző részben az  $(\mathcal{M}_n^+, K^{(n),f})$  komplex állapotér skalárgömbületét határoztuk meg az  $n = 2$  esetben, ismertebb  $f$  függvények mellett, és igazoltuk, hogy nem minden  $f \in \mathcal{F}_{[0,\infty]}^{(S,n)}$  függvény 2-elfogadható. A jelen részben főként numerikus módszerekkel megvizsgáljuk, hogy mely ismertebb metrikák lesznek 3- illetve 4-elfogadhatók. Az  $\mathcal{M}_3^+$  sokaság esetén már meglehetősen bonyolult kifejezések adódnak a skalárgömbületre. Az ilyen irányú vizsgálatokból eddig csak a skalárgömbület értékét lehetett ismerni a legkevésbé állapotban, mely egyszerűen adódik a (4.80) és (4.81) képletből.

Először az  $n = 3$  esetben vizsgáljuk meg az ismertebb metrikák skalárgömbületét komplex állapotéren.

A legkisebb metrika esetén az  $x, y, z$  sajátértékekkel rendelkező  $D \in \mathcal{M}_3^+$  állapotban az  $(\mathcal{M}_3^+, K_{SM}^{(3)})$  Riemann-sokaság skalárgömbülete

$$\text{Scal}_{SM}(x, y, z) = \frac{3}{2} \frac{3 + 7(xy + yz + zx)}{(x + y)(y + z)(z + x)} + 7. \quad (4.290)$$

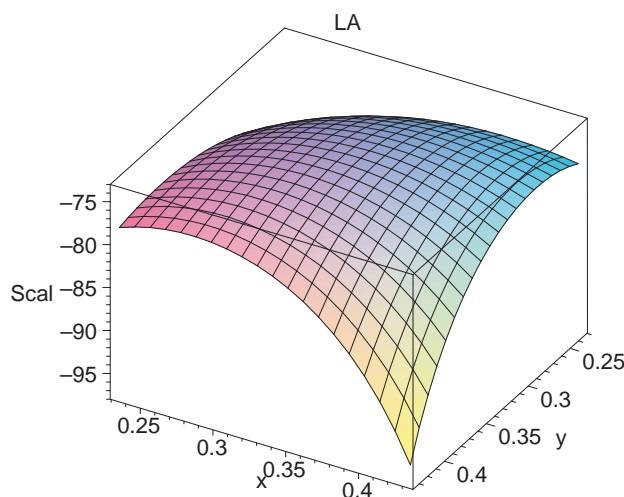
Matematikailag igazolható, hogy az esetben a skalárgömbületnek lokális minimuma van a legkevésbé állapotban, mint ahogy azt az alábbi ábra mutatja. (Az ábrán a skalárgömbület mint az  $x, y$  független sajátértékek függvénye szerepel, ugyanis ekkor a  $x + y + z = 1$  feltétel egyértelműen meghatározza  $z$ -t.)



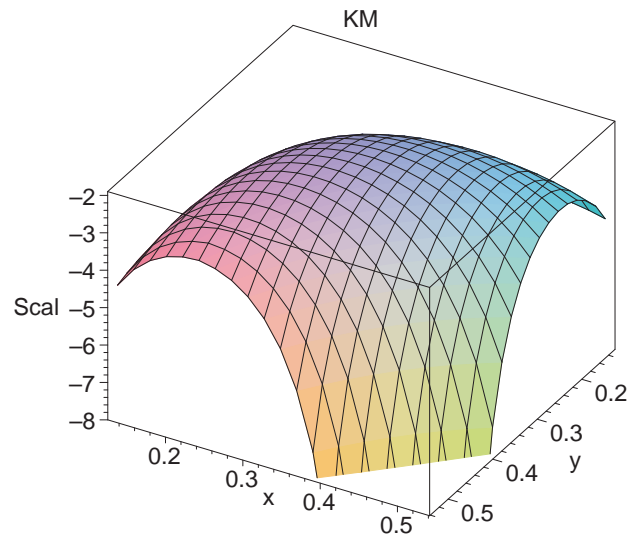
A *legnagyobb metrika* esetén a skalárgörbületet az alábbi kifejezés adja meg.

$$\begin{aligned} \text{Scal}_{\text{LA}}(x, y, z) = & -\frac{1}{2} \frac{31(x^2 + y^2 + z^2) + 85(xy + yz + zx)}{(x+y)(y+z)(z+x)} - \\ & - 7 \frac{(x^2y^2(x+y) + y^2z^2(y+z) + z^2x^2(z+x))}{xyz(x+y)(y+z)(z+x)} + 7 \end{aligned} \quad (4.291)$$

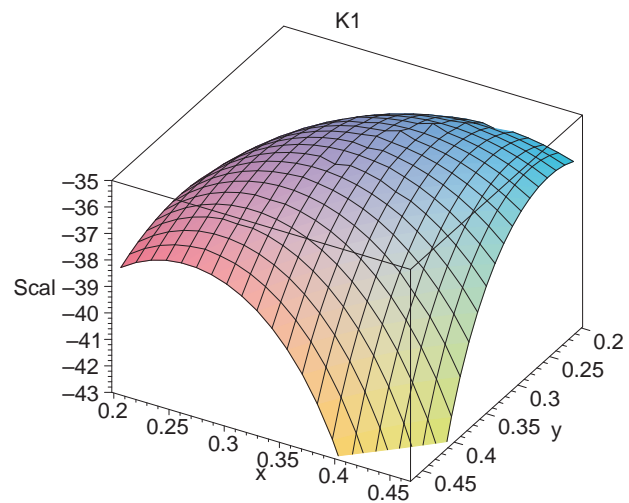
Ekkor a skalárgörbület lokális maximummal rendelkezik a legkevertebb állapotban.



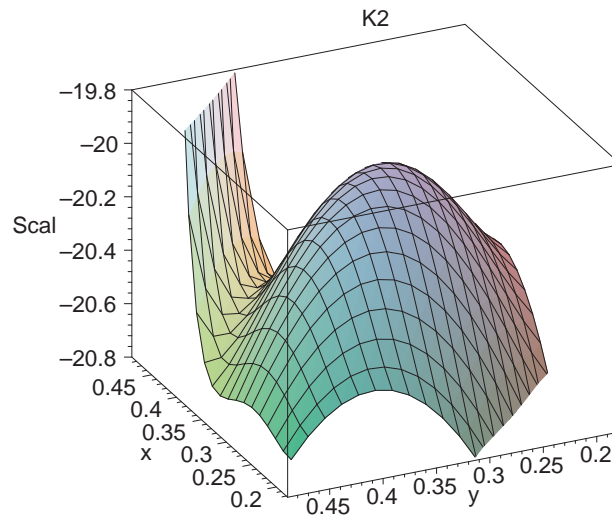
A *Kubo–Mori-metrika* esetén a (4.80) képlettel a skalárgörbületre adódó kifejezést nem lehet tovább egyszerűsíteni, ezért most csak ábrán mutatjuk be a viselkedését.



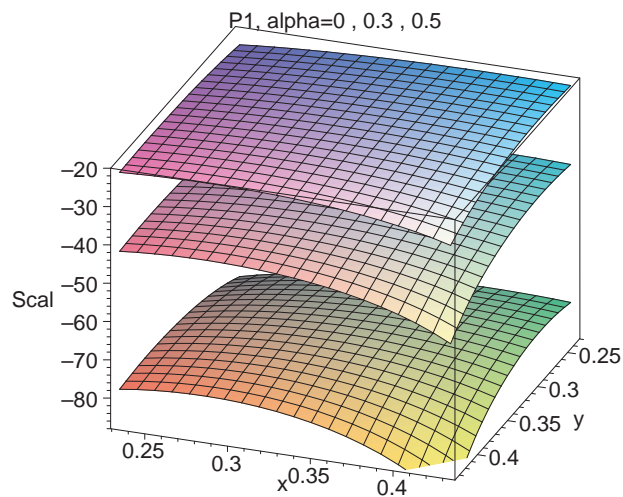
A *K1* és *K2* metrikákat vizsgálva kiderül, hogy a mindkét metrikának lokális maximuma van a legkevertebb állapotban.



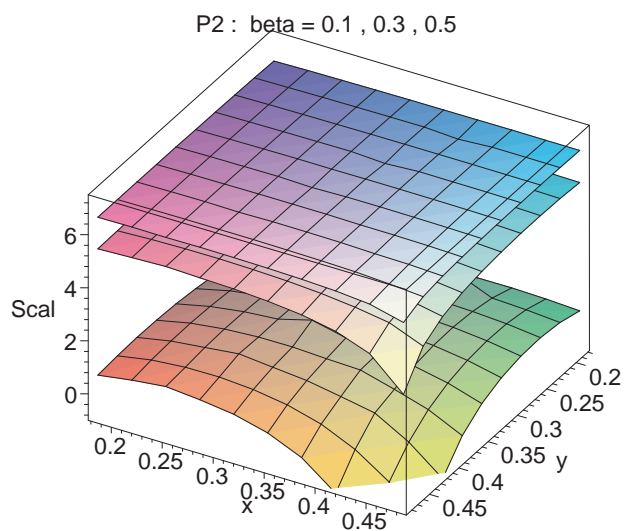
Ez a lokális maximum azonban nem globális a K2 metrika esetén a következő ábra szerint.



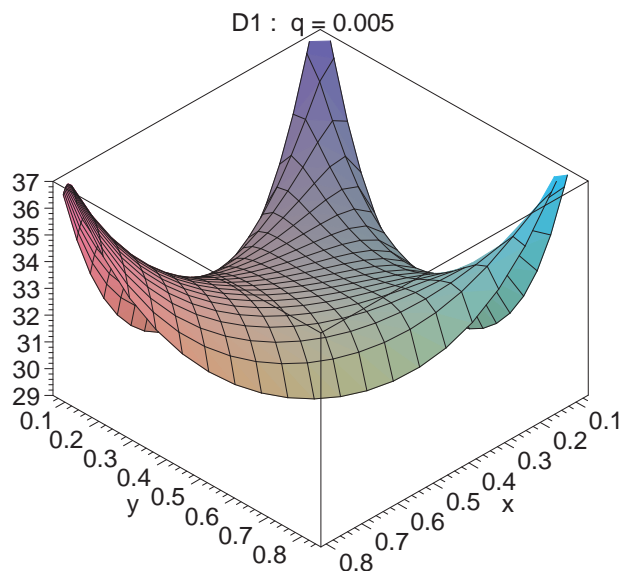
A  $P1$  és  $P2$  metrikacsaládnál a  $P1$  esetben adott állapotban a skalárgörbület monoton csökken az  $\alpha$  paraméter függvényében numerikus szimulációk alapján.



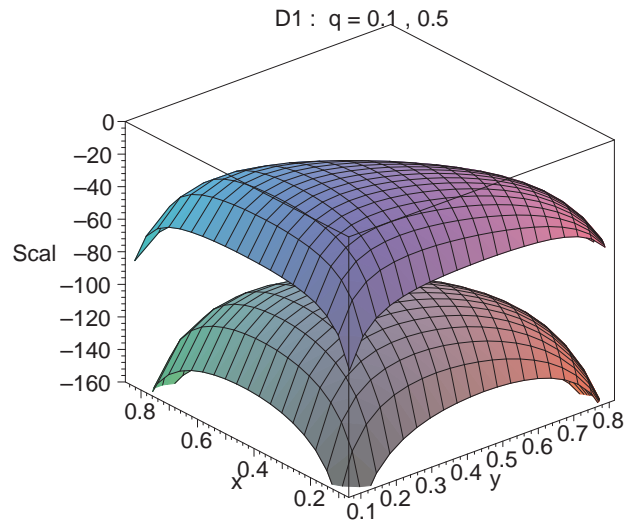
A P2 esetben az adott állapotbeli skalárgörbület nem monoton csökkenő függvénye a  $\beta$  paraméternek; az alábbi ábrán azonban a legnagyobb skalárgörbület-függvény tartozik a legkisebb  $\beta$  paraméterhez, és a legkisebb függvény pedig a legnagyobb  $\beta$  paraméterhez.



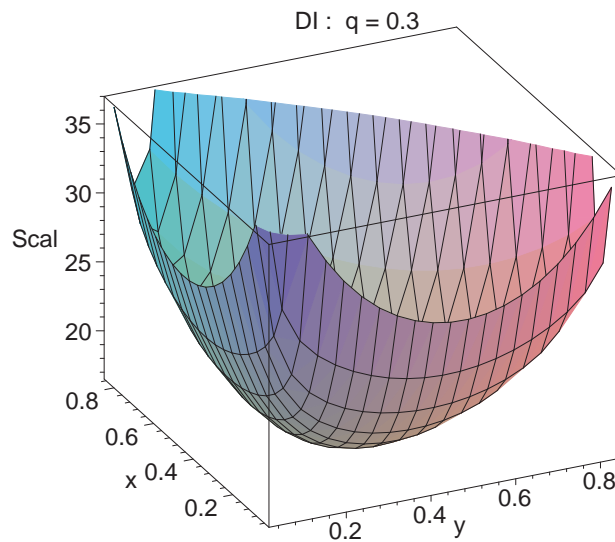
A *D1 metrika* alaposabb vizsgálata során kiderül, hogy nagyon kis  $q$  paraméter esetén a skalárgörbületnek sem globális minimuma, sem maximuma nincs a legkevertebb állapotban.



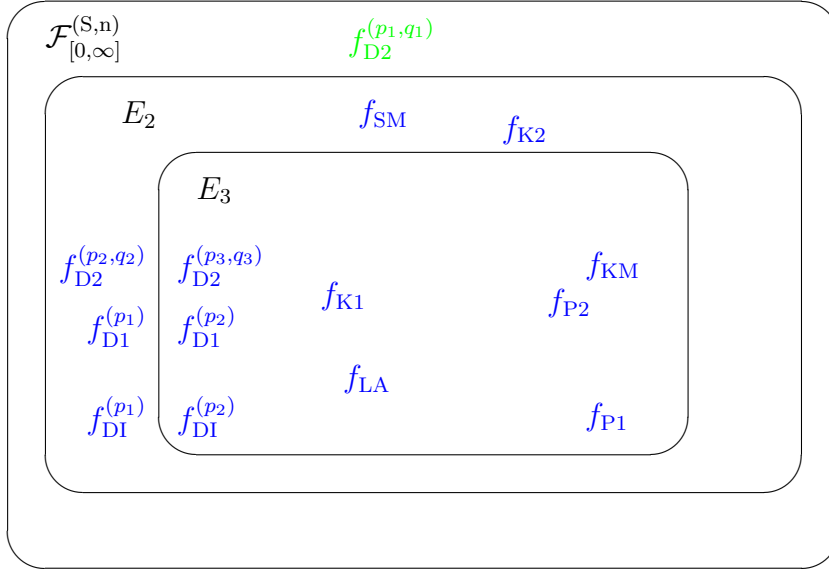
Nagyobb  $q$  paraméter esetén, azonban a megszokott módon viselkedik a skalárgörbület.



A  $DI$  metrikához tartozó skalárgörbületnek mind lokális minimuma, mind maximuma lehet a legkevertebb állapotban, most az előzőre mutatunk példát.



Tehát a metrikák 3-elfogadhatóságát vizsgálva, az alábbi ábrán foglalhatjuk össze eredményeinket.



Az  $\mathcal{M}_4^+$  sokaság esetén a skalárgömbületre adódó kifejezés annyira bonyolult, hogy numerikus vizsgálata sem egyszerű, különösen érvényes ez a paraméteres metrikacsaládok (például  $f_{P1}$ ,  $f_{P2}$ ,  $f_{D1}$ ) esetére. Ezért csak az  $f_{LA}$ ,  $f_{KM}$  és  $f_{K1}$  függvények által generált metrika 4-elfogadhatóságát vizsgáljuk numerikusan.

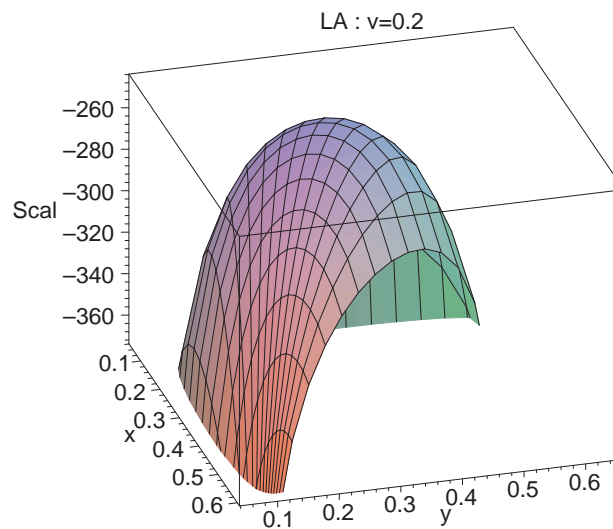
A *LA metrika esetén* a skalárgömbület az  $(\mathcal{M}_4^+, K_{LA}^{(4)})$  Riemann-sokaság  $x_1, x_2, x_3, x_4$  sajátértékkel rendelkező  $D \in \mathcal{M}_4^+$  állapotában a (4.80) formula alapján az

$$\text{Scal}(D) = -\frac{5}{2} \sum_{i=1}^n x_i \left( \sum_{j=1}^n \frac{1}{x_i + x_j} \right)^2 + 4 \sum_{i,j=1}^4 \frac{1}{x_i + x_j} - \frac{119}{8} \sum_{i=1}^n \frac{1}{x_i} + \frac{105}{4} \quad (4.292)$$

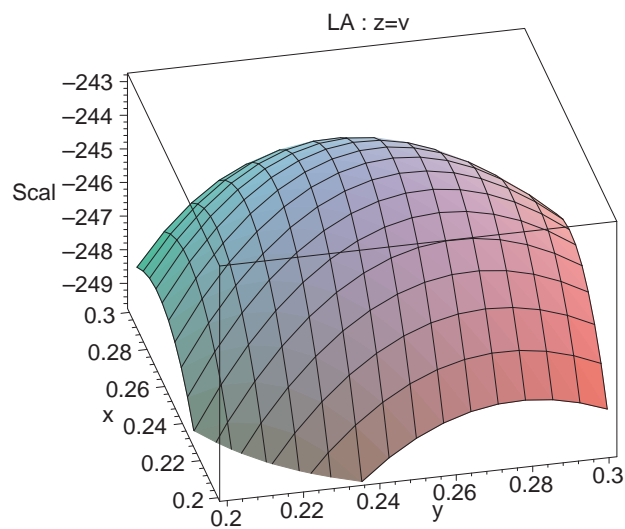
alakra hozható. Ha a  $D$  állapot sajátértékei  $x, y, z$  és  $v$ , akkor  $D$  állapotot gyakran a  $D_{x,y,z,v}$  formában fogjuk írni ebben a részben.

Rögzített  $v$  érték esetén, a skalárgömbületet az  $x$  és  $y$  változók függvényének tekintve numerikus szimulációk alapján azt kapjuk, hogy a skalárgömbület konkáv függvény, mely legnagyobb értékét a legkevertebb állapotban veszi fel. Példaként a  $v = 0.2$  érték esetén ábrázoljuk a skalárgömbület függvényt.





Valamivel egyszerűbb kifejezést kapunk a skalárgörbületre, ha feltételezzük, hogy valamelyik sajátérték multiplicitása kettő, vagyis például  $z = v$  teljesül. Az  $x, y, z, z$  sajátértékekkel rendelkező állapotokban a skalárgörbület értékét az alábbi ábra szemlélteti az  $x$  és  $y$  változó függvényeként

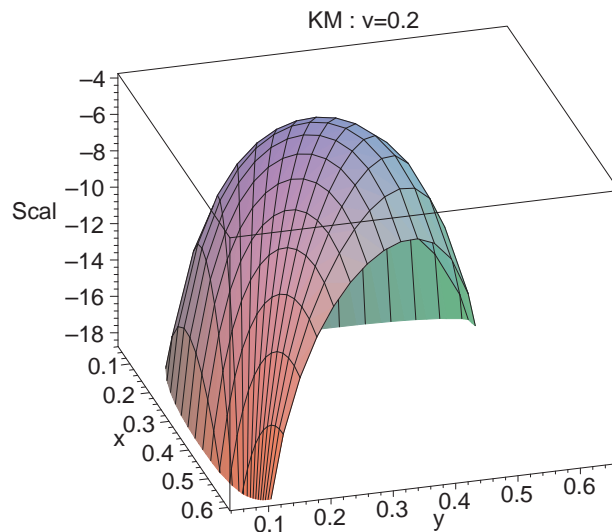


Amennyiben további sajátértékek is megegyeznek, matematikailag egyszerűen bizonyíthatóvá válik, hogy a skalárgörbület szigorúan konkáv függvény, mely legnagyobb értékét a legkevertebb állapotban veszi fel. Ekkor az alábbi, könnyen kezelhető formulákat kapjuk

$$\text{Scal}(D_{x,y,y,y}) = -\frac{1}{2} \frac{85x^3 + 258x^2y + 168y^2x + 29y^3}{xy(x+y)^2} + \frac{105}{4} \quad (4.293)$$

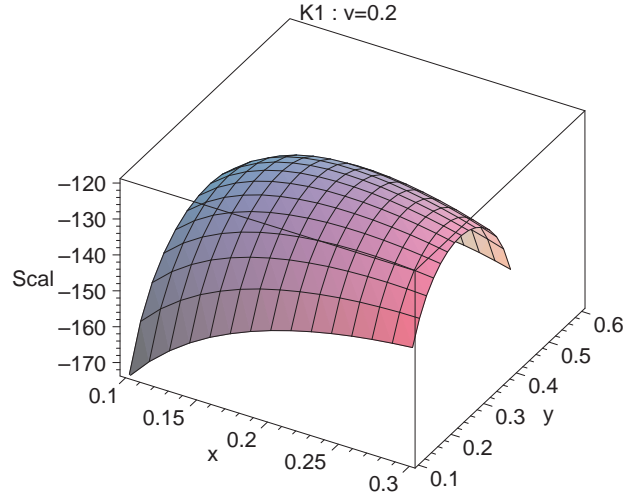
$$\text{Scal}(D_{x,x,y,y}) = -\frac{1}{4} \frac{99x^3 + 433x^2z + 433y^2x + 115y^3}{xy(x+y)^2} + \frac{105}{4} .$$

A *KM metrika esetén* a skalárgörbületre meglehetősen bonyolult formula adódik a (4.80) képlet alapján. A numerikus számítások azt mutatják, hogy a skalárgörbület konkáv függvény, mely legnagyobb értékét a legkevertebb állapotban veszi fel. Példaként a  $v = 0.2$  érték esetén ábrázoljuk a skalárgörbület-függvényt.



A sajátértékek további megegyezése esetén sem kapunk könnyen kezelhető formulát a skalárgörbületre, azonban a numerikus szimulációk ezekben az esetekben is megerősítik Petz sejtését.

A *K1 metrika esetén* a skalárgörbület szintén nagyon bonyolult függvény lesz. A numerikus szimulációk azonban erre a függvényre is azt mutatják, hogy szigorúan konkáv, és a legnagyobb értékét a legkevertebb állapotban veszi fel. Példaként a  $v = 0.2$  érték esetén ábrázoljuk a skalárgörbület függvényt.



Az elvégzett numerikus szimulációk alátámasztják, hogy  $f_{LA}, f_{KM}, f_{K1} \in E_4$ .

#### 4.5. A térfogat Taylor-sorfejtése

Az  $(\mathcal{M}_2^+, K^{(2),f})$  Riemann-sokaság több differenciálgeometriai jellemzőjét meg lehet határozni, ennek az egyik oka a sokaság alacsony dimenziószáma, a másik pedig, hogy ezen a sokaságon minden  $K^{(2),f}$  monoton metrika felírható

$$K_{11}^{(2),f} = \frac{1}{1-r^2} \quad (4.294)$$

$$K_{22}^{(2),f} = \frac{r^2}{(1+r)f\left(\frac{1-r}{1+r}\right)}$$

$$K_{33}^{(2),f} = \frac{r^2 \sin^2 \theta}{(1+r)f\left(\frac{1-r}{1+r}\right)}$$

alakban, ahol az  $\mathcal{M}_2^+$  sokaság gömbi paraméterezését használtuk, és a koordináták  $(r, \vartheta, \varphi)$  sorrendjét rögzítettük.

Míg a klasszikus esetben a  $\mathcal{P}_n$  tér térfogata egy jól meghatározott véges mennyiség volt (lásd a 2.9. tételt), az  $\mathcal{M}_2^+$  sokaság térfogatát alapvetően befolyásolja az  $f \in \mathcal{F}_{[0,\infty]}^{(S,n)}$  függvény. A sokaság térfogatát az invariáns sűrűségfüggvénynek a sokaságon vett integráljaként kapjuk a jelen esetben

$$V = \int_0^1 \int_0^\pi \int_0^{2\pi} \frac{r^2 \sin \vartheta}{(1+r)\sqrt{1-r^2}f\left(\frac{1-r}{1+r}\right)} d\varphi d\vartheta dr . \quad (4.295)$$

A  $t = \frac{1-r}{1+r}$  változócsere és két integrál kiszámolása után a

$$V = 2\pi \int_0^1 \left( \frac{1-t}{1+t} \right)^2 \frac{1}{\sqrt{t}f(t)} dt \quad (4.296)$$

kifejezés adódik. Ezen formula alapján az  $\mathcal{M}_2^+$  sokaság térfogata az alábbiak szerint alakul az  $f \in \mathcal{F}_{[0,\infty]}^{(S,n)}$  függvénytől függően.

$$f_{\text{SM}}(x) : \quad V_{\text{SM}} = \pi^2 \quad (4.297)$$

$$f_{\text{LA}}(x) : \quad V_{\text{LA}} = \infty$$

$$f_{\text{KM}}(x) : \quad V_{\text{KM}} = 2\pi^2$$

$$f_{\text{K1}}(x) : \quad V_{\text{K1}} = 8\pi$$

$$f_{\text{D1}}(x) : \quad V_{\text{D1}} = \infty$$

$$f_{\text{DI}}(x) : \quad V_{\text{DI}} = 2\pi \frac{\arctan \sqrt{\frac{q}{1-q}} + \arctan \frac{1-q}{\sqrt{q-q^2}} - \pi \sqrt{q-q^2}}{(1-2q)^2 \sqrt{q-q^2}}$$

$$f_{\text{WY}}(x) : \quad V_{\text{WY}} = 4\pi(\pi - 2)$$

(A többi speciális operátormonoton függvényre az integrálás nehezen végezhető el, illetve csak egyéb transzcendens függvényekre vezethető vissza.)

Adott  $0 < c < 1$  paraméter esetén jelölje  $S_2^{(c)}$  a háromdimenziós euklideszi térben lévő origó körüli  $c$  sugarú gömbfelszín. Az

$$i : S_2^{(c)} \rightarrow \mathcal{M}_2^+ \quad (c, \vartheta, \varphi) \mapsto (c, \vartheta, \varphi) \quad (4.298)$$

identikus leképezés beágyazás, vagyis az  $S_2^{(c)}$  halmaz részsokasága az  $\mathcal{M}_2^+$  térnek. Az  $i$  beágyazás által indukált  $i^*K^{(2),f}$  Riemann-metrika az  $S_2^{(c)}$  sokaságon a

$$(i^*K^{(2),f})_{11} = \frac{r^2}{(1+r)f\left(\frac{1-r}{1+r}\right)} \quad (4.299)$$

$$(i^*K^{(2),f})_{22} = \frac{r^2 \sin^2 \theta}{(1+r)f\left(\frac{1-r}{1+r}\right)}$$

alakban írható fel, ahol a  $(\vartheta, \varphi)$  koordinátázást használjuk az  $S_2^{(c)}$  sokaságon. Megállapíthatjuk, hogy az  $i^*K^{(2),f}$  metrika megegyezik a háromdimenziós euklideszi tér által indukált Riemann-metrika pozitív számszorosásával az  $S_2^{(c)}$  sokaságon. Az  $S_2^{(c)}$  halmazon a geodetikuskok a főkörök.

Az  $S_2^{(c)}$  részsokaságon a metrika megegyezik a gömbfelszín metrikájának számszorosával, ezért szimmetria okokból az  $(\mathcal{M}_2^+, K^{(2),f})$  tér geodetikusainak a vizsgálatát elég a  $\vartheta = \frac{\pi}{2}$  síkon elvégezni.

**4.15. Tétel.** Az  $(\mathcal{M}_2^+, K^{(2),f})$  Riemann-sokaság esetén a

$$\gamma : \mathbb{R} \rightarrow \mathcal{M}_2^+ \quad t \mapsto \gamma(t) = \left( r(t), \frac{\pi}{2}, \varphi(t) \right) \quad (4.300)$$

görbe pontosan akkor geodetikus, ha

$$\frac{d^2 r(t)}{dt^2} + \frac{r}{1-r^2} \left[ \frac{dr(t)}{dt} \right]^2 - \frac{r(1-r) \left( (2+3r+r^2)f\left(\frac{1-r}{1+r}\right) + 2rf'\left(\frac{1-r}{1+r}\right) \right)}{2(1+r)^2 f^2\left(\frac{1-r}{1+r}\right)} \left[ \frac{d\varphi(t)}{dt} \right]^2 = 0 \quad (4.301)$$

$$\frac{d^2 \varphi(t)}{dt^2} + \frac{(2+3r+r^2)f\left(\frac{1-r}{1+r}\right) + 2rf'\left(\frac{1-r}{1+r}\right)}{r(1+r)^2 f^2\left(\frac{1-r}{1+r}\right)} \frac{d\varphi(t)}{dt} \frac{dr(t)}{dt} = 0$$

teljesül.

A tételben szereplő differenciálegyenletrendszer megoldása tetszőleges  $f \in \mathcal{F}_{[0,\infty]}^{(S,n)}$  függvény esetén nem ismert. A sugárirányú geodetikusokat azonban meg lehet határozni. A tételben szereplő egyenletrendszernek ugyanis egy megoldása a

$$\gamma : \left[ 0, \frac{\pi}{2} \right[ \rightarrow \mathcal{M}_2^+ \quad t \mapsto \gamma(t) = (\sin t, \vartheta_0, \varphi_0) \quad (4.302)$$

függvény, alkalmas  $\vartheta_0$  és  $\varphi_0$  paraméter esetén. Ezek alapján az  $(\mathcal{M}_2^+, K^{(2),f})$  térben a legkevertebb állapot körüli  $R$  ( $0 < R < \frac{\pi}{2}$ ) sugarú nyílt gömböt a

$$S_{\tilde{D}}(R) = \{ (r, \vartheta, \varphi) \in \mathcal{M}_2^+ \mid 0 \leq r < \sin R \} \quad (4.303)$$

halmaz adja meg, ahol  $\tilde{D}$  jelöli a legkevertebb állapotot. Ez a halmaz az euklideszi térben lévő  $\sin R$  sugarú nyílt gömb.

**4.16. Tétel.** Az  $(\mathcal{M}_2^+, K^{(2),f})$  Riemann-sokaság  $\tilde{D}$  legkevertebb állapota körüli  $R$  ( $0 < R < \frac{\pi}{2}$ ) sugarú gömb térfogata

$$V(S_{\tilde{D}}(R)) = 4\pi \int_0^R \frac{r^2}{(1+r)\sqrt{1-r^2}f\left(\frac{1-r}{1+r}\right)} dr, \quad (4.304)$$

melynek  $R$  szerinti sorfejtése az alábbi.

$$\begin{aligned}
 V(R) = \frac{4\pi R^3}{3} & \left[ 1 - \frac{1 + 6f''(1)}{5} R^2 + \frac{2 - 30f^{(4)}(1) + 150f''(1) + 180(f''(1))^2}{105} R^4 + \right. \\
 & + \left( -\frac{1}{945} + \frac{14}{9}f^{(4)}(1) + \frac{178}{15}f''(1) - 4(f''(1))^2 - \frac{4}{135}f^{(6)}(1) + \right. \\
 & \left. \left. + \frac{32}{3}f^{(3)}(1) + \frac{8}{9}f''(1)f^{(4)}(1) - \frac{8}{3}(f''(1))^3 \right) R^6 + O(R^8) \right]
 \end{aligned} \tag{4.305}$$

A tételben szereplő sorfejtés és a 2.14. tétel szerint az  $(\mathcal{M}_2^+, K^{(2),f})$  sokaság skalárgömbülete a legkevertebb állapotban

$$\text{Scal}(\tilde{D}) = 6 + 36f''(1), \tag{4.306}$$

összhangban a 4.12. tétellel.

Az  $(\mathcal{M}_2^+, K^{(2),f})$  sokaságban a  $D \neq \tilde{D}$  állapot körüli  $R$  sugarú gömb meghatározásához ismerni kellene a 2.17. definícióval értelmezett  $D$  ponthoz tartozó exponenciális leképezést. Ez azonban általában nem lehetséges, mert a  $D$  pontból induló geodetikusokat nem ismerjük.

**4.3. Példa.** Tekintsük az  $(\mathcal{M}_2^+, K^{(2),f})$  Riemann-sokaság

$$D = \begin{pmatrix} \frac{1}{2} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{2} \end{pmatrix} \in \mathcal{M}_2^+ \tag{4.307}$$

pontját. Ezt a pontot az  $(r_0, \frac{\pi}{2}, 0)$  paraméterekkel jellemezhetjük, ahol  $r_0 = \frac{1}{2}$ . A  $v = (a, b, 0) \in T_D \mathcal{M}_2^+$  érintővektorra ekkor

$$K_D^{(2),f}(v, v) = \frac{a^2}{1 - r_0^2} + \frac{b^2 r^2}{(1 + r_0)f\left(\frac{1-r_0}{1+r_0}\right)} \tag{4.308}$$

teljesül.

Minden  $\alpha \in [0, 2\pi[$  paraméterre jelölje  $\gamma_\alpha$  azt a geodetikust, melyre  $\gamma_\alpha(0) = D$  és  $\dot{\gamma}_\alpha(0) = (a, b, 0)$  teljesül, ahol

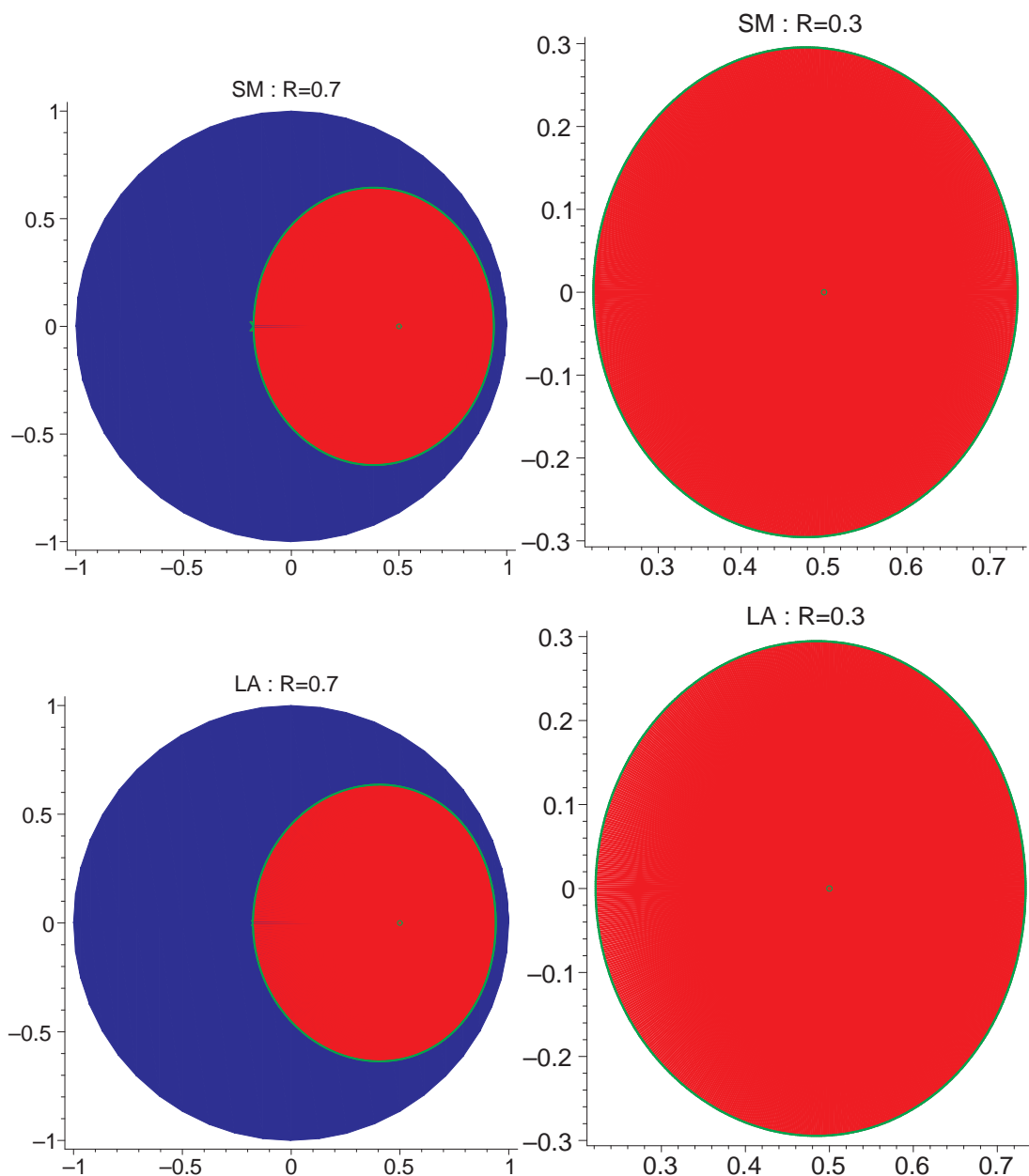
$$a = \sqrt{1 - r_0^2} \sin \alpha \quad b = \sqrt{(1 + r_0)f\left(\frac{1 - r_0}{1 + r_0}\right)} \frac{\cos \alpha}{r_0}. \tag{4.309}$$

Ekkor megfelelően kicsi  $R$  paraméterek esetén a

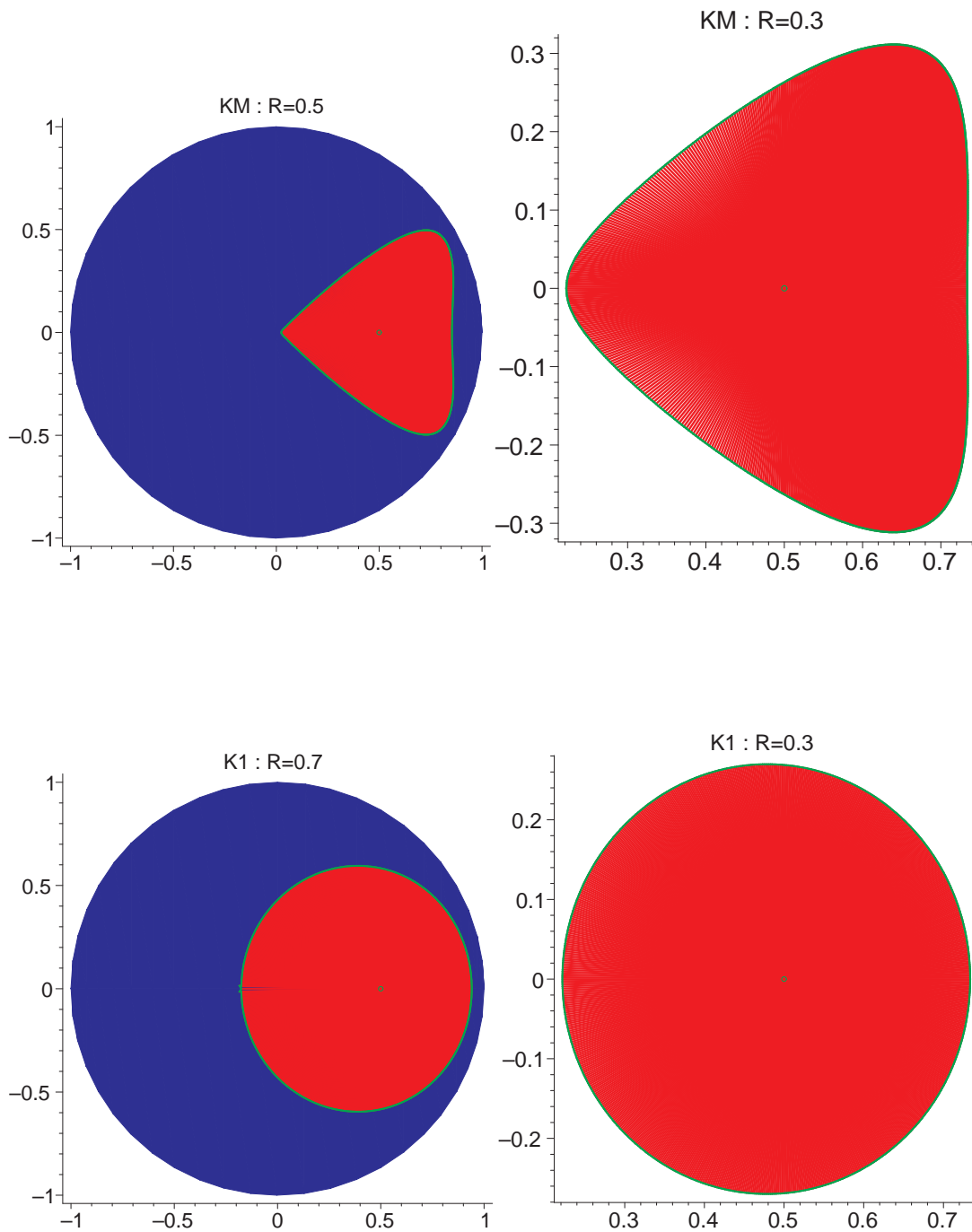
$$S_D(R) = \{\gamma_\alpha(t) \mid \alpha \in [0, 2\pi[, 0 \leq t < R\} \tag{4.310}$$

tartomány a  $D$  állapot körüli  $R$  sugarú nyílt gömb lesz.

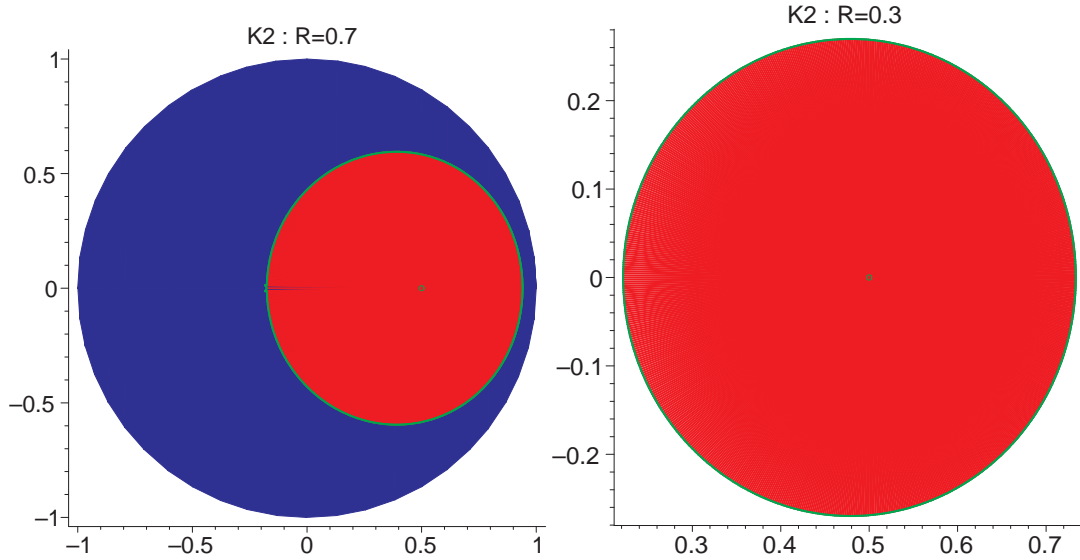
Az alábbi ábrákon látható, hogy az  $f_{SM}, f_{LA}, f_{KM}, f_{K1}, f_{K2}$  függvények esetén hogy néz ki a  $D$  pont körüli  $R = 0.7$  illetve  $R = 0.3$  sugarú nyílt gömb  $\vartheta_0 = \frac{\pi}{2}$  síkbeli metszete, numerikus szimulációk alapján. (A kék háttér az állapottér kiterjedését mutatja az  $x, y$  síkon, a piros tartomány pedig a zöld pont körüli  $R$  sugarú nyílt gömböt.)



Figyelem: a Kubo–Mori-metrikánál az egyik sugár nem 0.7, hanem 0.5 !







Az imént definiált, illetve az ábrákon is bemutatott  $S_D(R)$  tartományok térfogatának a meghatározása nem ismert általános  $f \in \mathcal{F}_{[0,\infty]}^{(S,n)}$  függvényre.

□

Az  $(\mathcal{M}_2^+, K^{(2),f})$  Riemann-sokaság  $D \in \mathcal{M}_2^+$  pontja körüli  $R$  sugarú nyílt gömb térfogatának  $R$  szerinti sorfejtését meghatározhatjuk a 2.14. tétel segítségével. Általános  $K^{(2),f}$  monoton metrika esetén a 2.3. részben bevezetett differenciálgeometriai segédmenyiségeket meglehetősen nehéz meghatározni. (A Ricci-görcületet illetve a skalárgörcületet a 4.3. részben megadtuk általános  $f \in \mathcal{F}_{[0,\infty]}^{(S,n)}$  függvényre.)

A továbbiakban megmutatjuk, hogyan alakul a térfogat sugar szerinti sorfejtése speciális operátormonoton függvények esetén.

A *legkisebb metrika* esetén minden  $D \in \mathcal{M}_2^+$  állapotra ugyanúgy néz ki az állapot körüli  $R$  sugarú gömb térfogatának az  $R$  szerinti sorfejtése, nevezetesen

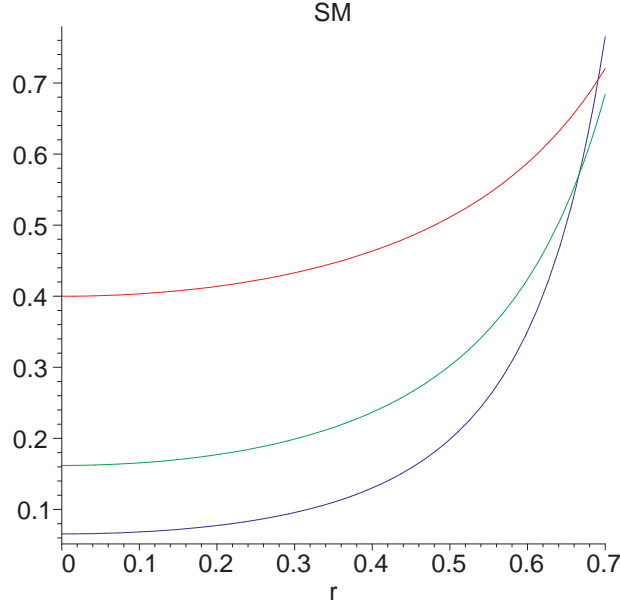
$$V(S_D(R)) = \frac{4\pi R^3}{3} \left( 1 - \frac{1}{5}R^2 + \frac{2}{105}R^4 - \frac{1}{945}R^6 + O(R^8) \right) \quad (4.311)$$

teljesül.

A *legnagyobb metrika* esetén tegyük fel, hogy a kiszemelt  $D \in \mathcal{M}_2^+$  állapot az  $S_2^{(r)}$  gömbhéjon fekszik. Ekkor

$$V(S_D(R)) = \frac{4\pi R^3}{3} \left( 1 - \frac{1}{15} \frac{6-r^2}{1-r^2} R^2 + \frac{1}{1575} \frac{3r^4 + 50r^2 + 225}{(1-r^2)^2} R^4 + \right. \quad (4.312) \\ \left. + \frac{1}{33075} \frac{r^6 - 84r^4 - 2380r^2 - 2170}{(1-r^2)^3} R^6 + O(R^8) \right)$$

teljesül. A sorfejtésben a zárójelben  $R^{2k}$  együtthatójaként megjelenő  $\alpha_k(r)$  függvényeket az alábbi ábra szemlélteti,  $k = 1, 2, 3$  esetre. Az  $\alpha_1, \alpha_2$  és  $\alpha_3$  függvény a piros, zöld és kék színnel van ábrázolva.



A térfogat sorfejtését a legkevertebb állapot körül a

$$V(S_{\bar{D}}(R)) = \frac{4\pi R^3}{3} \left( 1 + \frac{2}{5}R^2 + \frac{17}{105}R^4 + \frac{62}{945}R^6 + O(R^8) \right) \quad (4.313)$$

kifejezés adja meg.

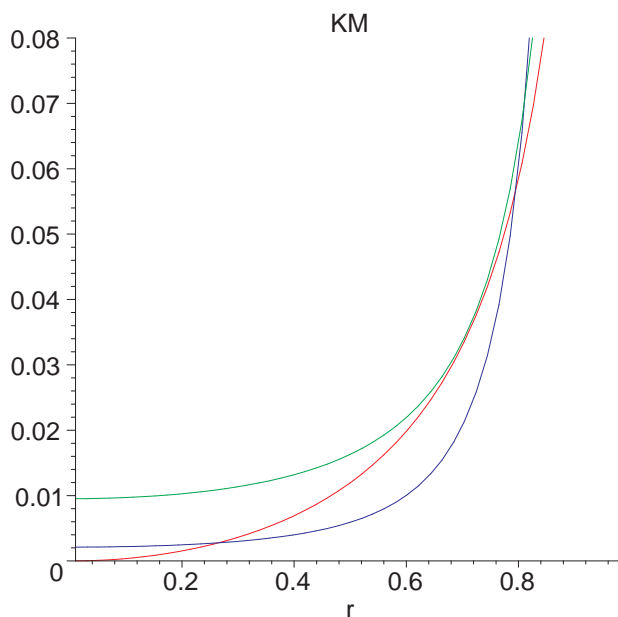
A *Kubo–Mori-metrika* esetén tetszőleges  $D \in \mathcal{M}_2^+$  állapot körül meghatározható a sorfejtés pontos alakja, azonban már bonyolultabb kifejezés adódik eredményül. A  $t = \frac{1-r}{1+r}$  változó segítségével

$$\alpha_1(r) = -\frac{4t(t^2 - t + 1) \log^2 t + 2(1 - t^4) \log t + (1 - t^2)^2}{60t(1 - t)^2 \log^2 t} \quad (4.314)$$

$$\begin{aligned} \alpha_2(r) = & \frac{1}{12600t^2(1 - t)^4 \log^4 t} \left( 8t^2(t^4 - 25t^3 - 22t^2 - 25t + 1) \log^4 t - \right. \\ & - (1 - t^2)(27t^6 - 62t^5 + 77t^4 + 172t^3 + 77t^2 - 62t + 27) \log^3 t - \\ & - 2(1 - t^2)^2(17t^4 - 23t^3 - 23t + 17) \log^2 t - \\ & \left. - 16(1 - t^2)^3(3t^2 - 2t + 3) \log t - 35(1 - t^2)^4 \right) \end{aligned} \quad (4.315)$$

$$\begin{aligned}
\alpha_3(r) = & -\frac{1}{8467200t^3(1-t)^6 \log^6 t} \left( 16t^3(64t^6 + 1507t^5 + 7020t^4 + \right. & (4.316) \\
& + 10594t^3 + 7020t^2 + 1507t + 64) \log^6 t + \\
& + 8(1-t^2)(225t^{10} - 774t^9 + 732t^8 + 472t^7 + 3691t^6 + \\
& + 8780t^5 + 3691t^4 + 472t^3 + 732t^2 - 774t + 225) \log^5 t + \\
& + (1-t^2)^2(2739t^8 - 8844t^7 + 8808t^6 - 5732t^5 - \\
& - 23126t^4 - 5732t^3 + 8808t^2 - 8844t + 2739) \log^4 t + \\
& + 8(1-t^2)^3(709t^6 - 2003t^5 + 1923t^4 + 246t^3 + 1923t^2 - 2003t + 709) \log^3 t + \\
& + 4(1-t^2)^4(2479t^4 - 5378t^3 + 3974t^2 - 5378t + 2479) \log^2 t + \\
& \left. + 56(1-t^2)^5(211t^2 - 266t + 211) \log t + 6944(1-t^2)^6 \right)
\end{aligned}$$

adódik. Az előző módszerhez hasonlóan ábrázoljuk a sorfejtésben szereplő  $\alpha_1, \alpha_2$  és  $\alpha_3$  függvényt.



A legkevertebb állapotban a térfogat sorfejtése

$$V(S_{\bar{D}}(R)) = \frac{4\pi R^3}{3} \left( 1 + \frac{1}{105}R^4 + \frac{2}{945}R^6 + O(R^8) \right) \quad (4.317)$$

lesz.

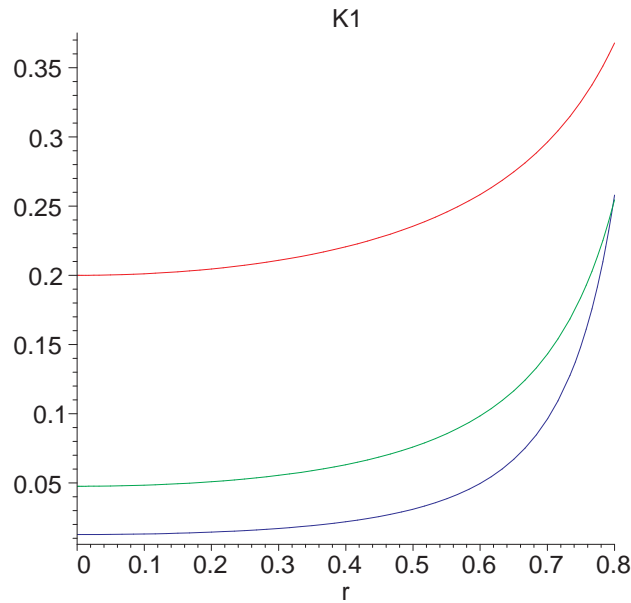
A *K1 metrika* esetén az  $(\alpha_i)_{i=1,2,3}$  függvényeket egyszerűbben ki tudjuk fejezni, ha bevezetjük a  $t = \frac{1-r}{1+r}$  jelölést. Ekkor

$$\alpha_1(r) = -\frac{(1-t)((1+t)\log t - (1-t))}{15t\log^2 t} \quad (4.318)$$

$$\begin{aligned} \alpha_2(r) = & -\frac{9(1-t^2)(3t^2+4t+3)\log^3 t + (1+t)^2(31t^2-26t+31)\log^2 t}{6300t^2\log^4 t} + \\ & + \frac{2(1-t^2)(13t^2+46t+13)\log t + 48(1-t)^2(t^2+t+1)}{6300t^2\log^4 t} \end{aligned} \quad (4.319)$$

$$\begin{aligned} \alpha_3(r) = & \frac{1}{2116800t^3\log^6 t} \left( -60(1-t^2)(15t^4+30t^3+32t^2+15)\log^5 t + \right. \\ & + 3(1+t)^2(457t^4-68t^3+82t^2-68t+457)\log^4 t + \\ & + 2(1-t^2)(781t^4+2286t^3+2674t^2+2286t+781)\log^3 t + \\ & + (1+t)^2(3137t^4+2320t^3+606t^2+2320t+3137)\log^2 t + \\ & + 288(1-t)(1+t)^3(17t^2+6t+17)\log t \\ & \left. + 1024(1-t)^2(t^2+t+1)(4t^2+7t+4) \right) \end{aligned} \quad (4.320)$$

teljesül. Az  $\alpha_1, \alpha_2$  és  $\alpha_3$  függvények viselkedése az alábbi ábrán látható.

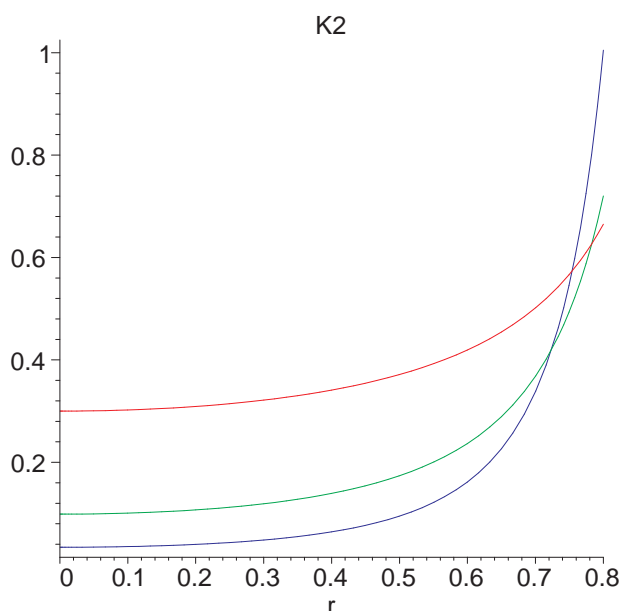


A legkevertebb állapotban a térfogat sorfejtése

$$V(S_{\bar{D}}(R)) = \frac{4\pi R^3}{3} \left( 1 + \frac{1}{5}R^2 + \frac{1}{21}R^4 + \frac{4}{315}R^6 + O(R^8) \right) \quad (4.321)$$

lesz.

A *K2 metrika* esetén az  $(\alpha_i)_{i=1,2,3}$  függvények az előző kifejezéseknél jóval bonyolultabbak lesznek, ezért csak az alábbi ábrán mutatjuk be viselkedésüket.

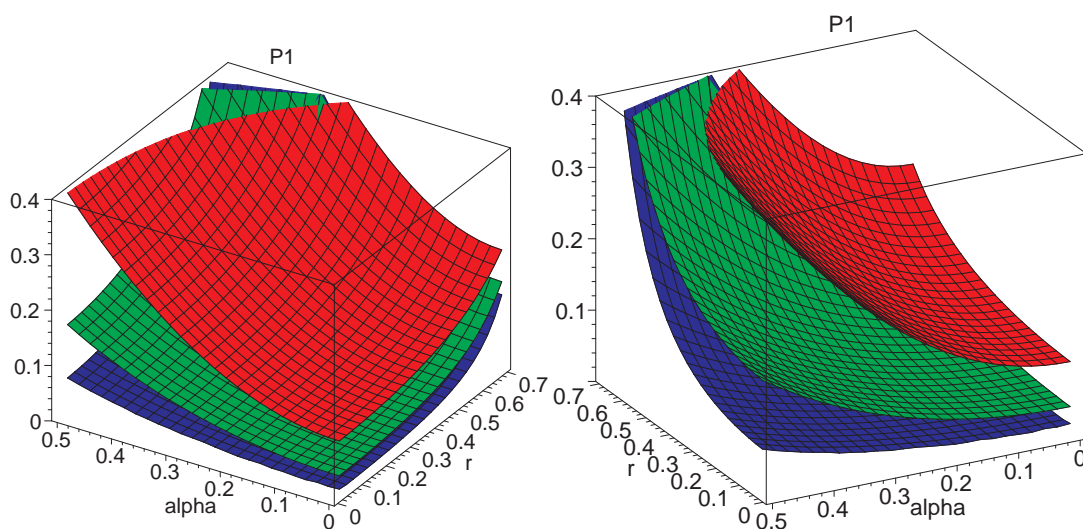


A legkevertebb állapotban a térfogat sorfejtése

$$V(S_{\bar{D}}(R)) = \frac{4\pi R^3}{3} \left( 1 + \frac{3}{10}R^2 + \frac{83}{840}R^4 + \frac{103}{3024}R^6 + O(R^8) \right) \quad (4.322)$$

lesz.

A *P1 metrikacsalád* esetén az  $(\alpha_i)_{i=1,2,3}$  függvények egyik változója az  $r$ , a másik pedig az  $\alpha$  paraméter. Ezen függvények bonyolultságára jellemző, hogy például az  $\alpha_3$  függvény „legoptimálisabb” alakjában is majdnem 500 tagot kell összeadni. Az előző ábrázolási módszereket meghagyva ( $\alpha_1$ –piros,  $\alpha_2$  zöld és  $\alpha_3$  kék) mutatjuk be az  $(\alpha_i)_{i=1,2,3}$  függvények viselkedését az  $r$  illetve az  $\alpha$  paramétertől függően.



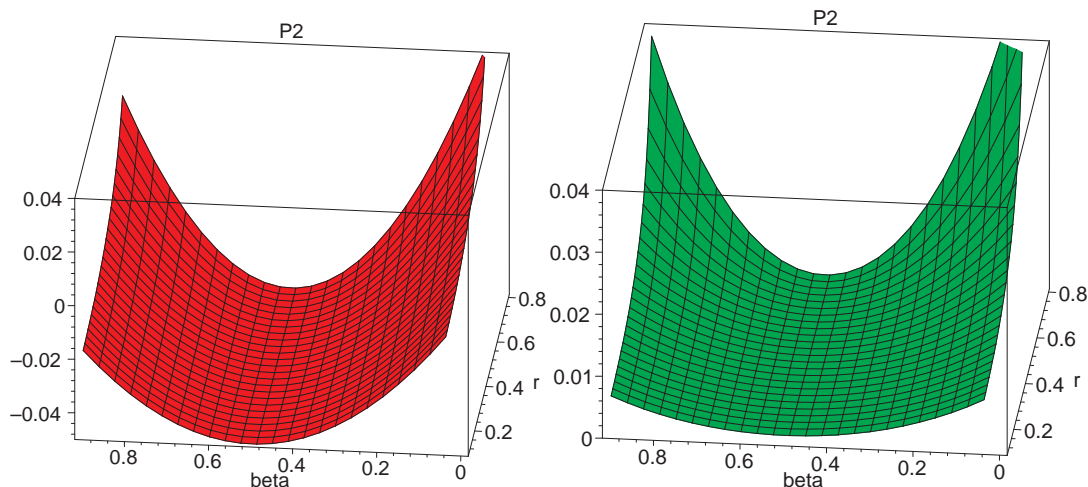
A legkevertebb állapotban a térfogat sorfejtése

$$V(S_{\bar{D}}(R)) = \frac{4\pi R^3}{3} \left( 1 + \frac{1 + 12\alpha^2}{10} R^2 + \frac{240\alpha^4 + 360\alpha^2 + 31}{840} R^4 + \right. \quad (4.323)$$

$$\left. + \frac{448\alpha^6 + 2800\alpha^4 + 2548\alpha^2 + 173}{15120} R^6 + O(R^8) \right)$$

lesz.

A  $P2$  metrikacsalád esetén az  $(\alpha_i)_{i=1,2,3}$  függvények egyik változója  $r$ , a másik pedig a  $\beta$  paraméter. Ezen esetben az  $\alpha_3$  függvénynek a „legoptimálisabb felírása” is 2363 segédváltozó bevezetését igényli, és 8798 tag összegeként állítható elő. Ilyen mértékű bonyolultság esetén azonban nagyon megnő a numerikus számításokból adódó hiba, ezért ezt a függvényt nem is ábrázoljuk. Az  $\alpha_1$  illetve  $\alpha_2$  függvény az alábbi ábrákon látható (piros, illetve zöld színnel).

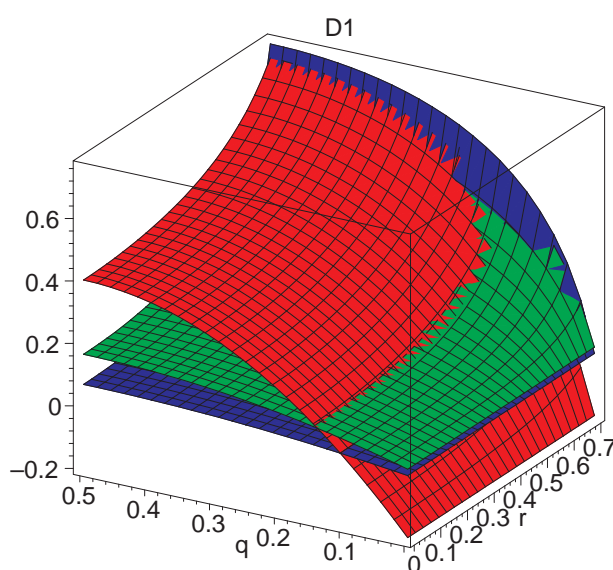


A legkevertebb állapotban a térfogat sorfejtése

$$V(S_{\bar{D}}(R)) = \frac{4\pi R^3}{3} \left( 1 + \frac{\beta^2 - \beta}{5} R^2 + \frac{2\beta^4 - 4\beta^3 + 6\beta^2 - 4\beta + 1}{105} R^4 + O(R^6) \right) \quad (4.324)$$

lesz.

A *D1 metrikacsalád* esetén az  $(\alpha_i)_{i=1,2,3}$  függvények egyik változója  $r$ , a másik pedig a  $q$  paraméter. Az  $(\alpha_i)_{i=1,2,3}$  függvények viselkedését az alábbi ábra mutatja.

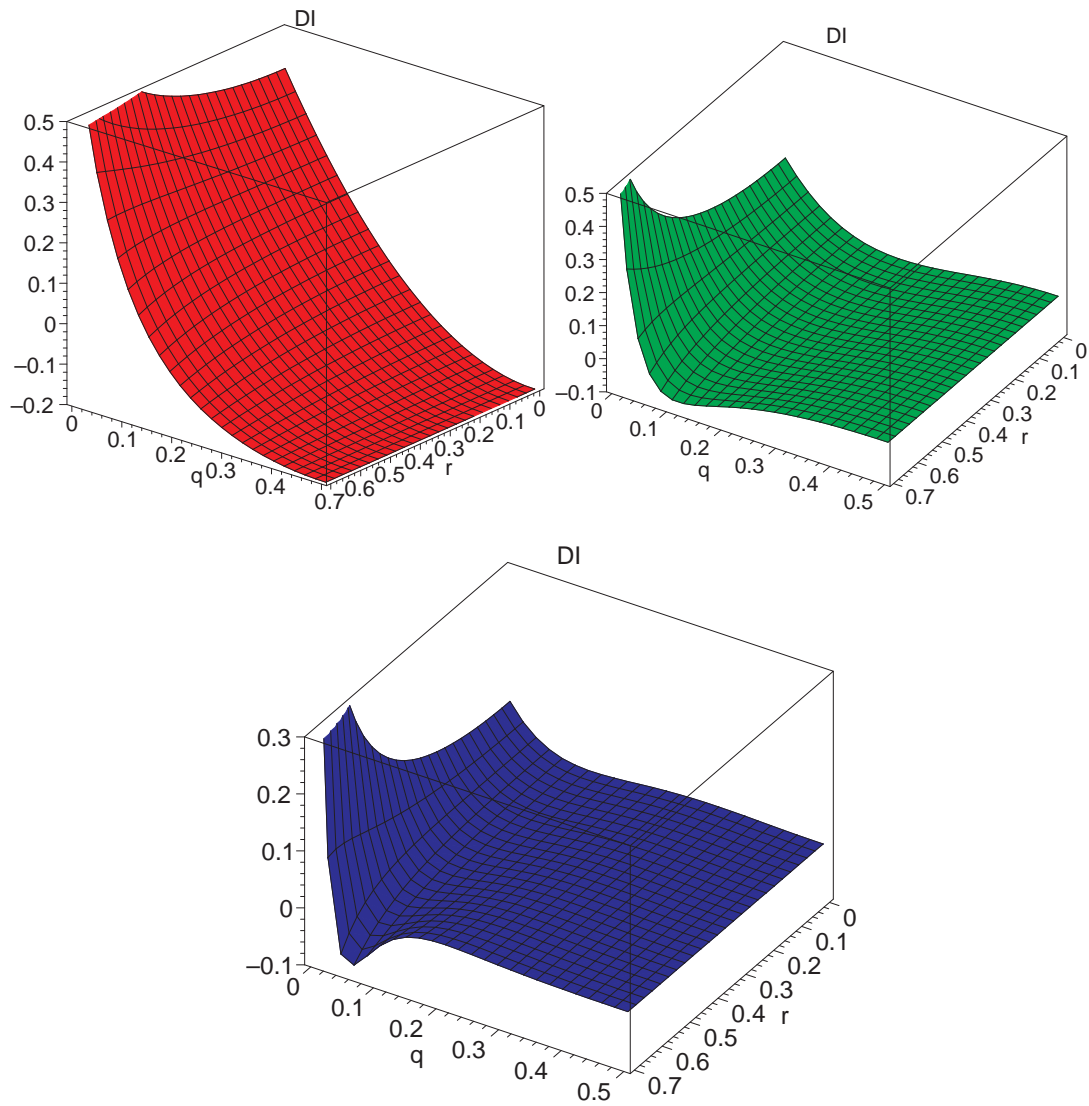


A legkevertebb állapotban a térfogat sorfejtése

$$V(S_{\bar{D}}(R)) = \frac{4\pi R^3}{3} \left( 1 - \frac{12q^2 - 12q + 1}{5} R^2 - \frac{60q^2 - 60q - 2}{105} R^4 - \frac{252q^2 - 252q + 1}{945} R^6 + O(R^8) \right) \quad (4.325)$$

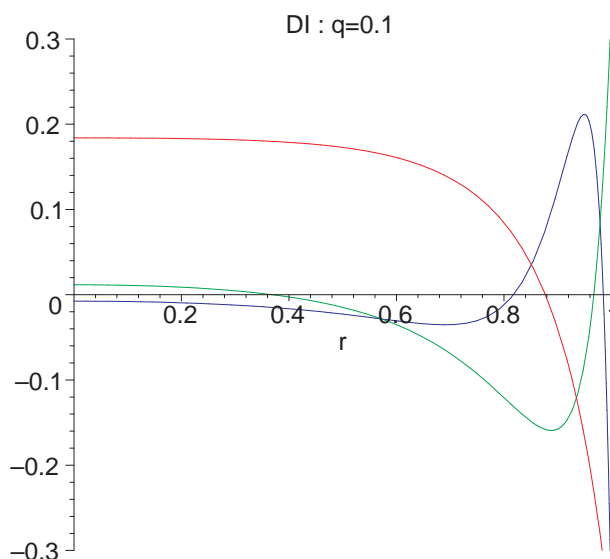
lesz.

A *DI metrikacsalád* esetén az  $(\alpha_i)_{i=1,2,3}$  függvények egyik változója az  $r$ , a másik pedig az  $q$  paraméter. Ebben az esetben, bizonyos  $q$  paraméterek esetén az  $(\alpha_i)_{i=1,2,3}$  függvények nem szigorúan monoton növekvő függvényei  $r$ -nek. A kétparaméteres  $(\alpha_i)_{i=1,2,3}$  függvények viselkedését az alábbi ábrák mutatják.





A  $q = \frac{1}{10}$  paraméter esetén az  $(\alpha_i)_{i=1,2,3}$  függvények nem szigorúan monoton növekednek az alábbi ábra szerint.



A legkevertebb állapotban a térfogat sorfejtése

$$V(S_{\bar{D}}(R)) = \frac{4\pi R^3}{3} \left( 1 + \frac{12q^2 - 12q + 2}{5} R^2 - \frac{720q^4 - 1440q^3 + 960q^2 - 240q + 17}{105} R^4 - \frac{20160q^6 - 60480q^5 + 70560q^4 - 40320q^3 + 11592q^2 - 1512q + 62}{945} R^6 + O(R^8) \right) \quad (4.326)$$

lesz.

Nagyobb  $n$  értékre meglehetősen nehéz a fenti vizsgálatok elvégzése, ugyanis a sorfejtésben szereplő mennyiségek kiszámításához olykor tízszeres szummát kell kiszámolni, ahol mindegyik futóindex egytől a tér dimenziószámáig megy. Így a műveletek száma legalább  $(n^2 - 1)^{10}$ . Becslések szerint az  $n = 2$ -ről áttérni az  $n = 3$  esetre kb 18.000-szer hosszabb gépidőt igényel.



## Jelölésgyűjtemény

| Jelölés:   | Rövid leírás:                          | Oldalszám: |
|--|--|------------|
| $\cong$  | függvények közötti ekvivalenciareláció | 108        |
| $\sim$   | függvények közötti ekvivalenciareláció | 115        |
| $\prec$  | majorálás relációja                    | 27, 89     |
| $[\cdot, \cdot]$   | vektormezők kommutátora                | 35         |
| $\langle \cdot, \cdot \rangle$   | Hilbert–Schmidt skalárszorzás          | 58, 94     |
| $\stackrel{\leq}{[0,1]}$   | függvények közötti reláció             | 101        |
| $\nabla$   | kovariáns deriválás                    | 37         |
| $\nabla^{(\alpha)}$  | $\alpha$ -kovariáns deriválás          | 38         |
| $\Gamma$   | Christoffel-szimbólum                  | 37         |
| $\kappa$   | Markov-féle magfüggvény                | 12         |
| $\sigma_1, \sigma_2, \sigma_3$   | Pauli-mátrixok                         | 88         |
| $\sigma_\mu^2$   | $\mu$ varianciája                      | 174        |
| $\sigma(H)$  | a $H$ mennyiség szórása                | 87         |
| $c_f$  | Cencov–Morozova-fél függvény           | 96         |
| $D_f$  | Csiszár-féle relatív entrópia          | 23         |
| $\left. \begin{array}{l} D_{\text{KL}}, D_{\text{H}}, D_{\chi^2}, \\ D_\alpha, D_{\text{B}}, D_{\text{Ha}}, \\ D_{\text{J}}, D_{\Delta}, D_{\text{LW}} \end{array} \right\}$ | relatív entrópiák                      | 22, 23     |
| $d_{\text{Bures}}(D_1, D_2)$   | állapotok Bures-távolsága              | 119        |
| $d_g(D_1, D_2)$  | állapotok $g$ -geodetikus távolsága    | 119        |
| $d(p, q)$  | a $p$ és $q$ eloszlás távolsága        | 76         |
| $E_\vartheta(\tilde{\vartheta})$   | várható érték                          | 13         |
| $E(H)$   | a $H$ mennyiség várható érték          | 87         |
| $E_n$  | elfogadható metrikák                   | 164        |

| Jelölés:   | Rövid leírás:                            | Oldalszám:   |
|--|--|--------------|
| $E_{n,\mu}$  | $\mu$ $n$ -edik momentuma                | 174          |
| $\exp_p$   | exponenciális leképezés                  | 44           |
| $\left. \begin{array}{l} \mathcal{F}_I, \mathcal{F}_I^{(n)}, \\ \mathcal{F}_I^{(S)}, \mathcal{F}_I^{(S,n)} \end{array} \right\}$   | operátormonoton függvények halmazai      | 95, 96       |
| $f^\setminus$  | az $f$ függvény transzponáltja           | 95, 113      |
| $f^\perp$  | az $f$ függvény duálisa                  | 95, 113      |
| $f^*g$   | a $g$ metrika visszahúzottja             | 50           |
| $f_*p$   | derivált leképezés                       | 50           |
| $\left. \begin{array}{l} f_{SM}, f_{LA}, f_{KM}, \\ f_{K1}, f_{K2}, f_{P1}, \\ f_{P2}, f_{WYD}, f_{WY}, \\ f_{P3}, f_{D1}^{(p)}, f_{D1}^{(q)}, \\ f_{D2}^{(p,q)} \end{array} \right\}$ | operátormonoton függvények               | 99, 178, 179 |
| $\left. \begin{array}{l} \mathcal{G}_I, \mathcal{G}_I^{(n)}, \\ \mathcal{G}_I^{(S)}, \mathcal{G}_I^{(S,n)} \end{array} \right\}$   | Radon-mértékek halmaza                   | 95, 96       |
| $g^{(F)}$  | Fisher-féle információs mátrix           | 9, 56        |
| $H(\cdot, \cdot)$  | relatív entrópia                         | 112          |
| $H_g(\cdot, \cdot)$  | $g$ -relatív entrópia                    | 113          |
| $H_g^{(szim)}$   | a $g$ -relatív entrópia szimmetrizáltja  | 114          |
| $H_{SM}, H_{(x-1)^2}, H_{\log}$  | relatív entrópiák                        | 117, 118     |
| $I_{D_{\vartheta_0}}^f$  | kvantummechanikai Fisher-információ      | 105          |
| $\left. \begin{array}{l} \mathcal{K}_I, \mathcal{K}_I^{(n)}, \\ \mathcal{K}_I^{(S)}, \mathcal{K}_I^{(S,n)} \end{array} \right\}$   | operátorkonvex függvények halmaza        | 113          |
| $K^{g,(n)}$  | a $g$ által generált metrika             | 115          |
| $K_D^{(n),f}$  | az $f$ által generált monoton metrika    | 95           |
| $L_{n,D}, R_{n,D}$   | balról illetve jobbról szorzás operátora | 94           |
| $l_\gamma(a, b)$   | a $\gamma$ görbe ívhossza                | 41           |
| $\mathcal{M}_n$  | az állapottér érintőtere                 | 93           |
| $\mathcal{M}_n^+$  | állapottér                               | 86           |
| $\tilde{\mathcal{M}}_n^+$  | az $\mathcal{M}_n^+$ tér bővítése        | 122          |
| $M_n^+$  | pozitív definit mátrixok                 | 55           |
| $\tilde{M}_n$  | az $M_n^+$ tér bővítése                  | 78           |

| Jelölés:  | Rövid leírás:                          | Oldalszám:   |
|---|--|--|
| $(M, \nabla)$   | differenciálgeometria                  | 37   |
| $(M, \mathcal{A})$  | differenciálható sokaság               | 32   |
| $(M, g)$  | Riemann-sokaság                        | 36   |
| $m_\mu$   | $\mu$ várható értéke                   | 174  |
| $m_{ij}, m_{ijk}$   | $m$ -mennyiségek                       | $\left\{ \begin{array}{l} 58, 61, \\ 100, 124 \end{array} \right.$ |
| $n : N \rightarrow TM$  | normálvektormező                       | 53   |
| $\mathcal{P}_n$   | diszkrét eloszlások halmaza            | 32   |
| $\left. \begin{array}{l} \ \text{Ric}\ ^2, \ R\ ^2, \dot{\text{Ric}}, \\ \Delta \text{Scal}, \ \nabla \text{Ric}\ ^2, \\ \langle \Delta \text{Ric}, \text{Ric} \rangle, \\ \ \nabla \text{Scal}\ ^2, \langle \nabla^2 \text{Scal}, \text{Ric} \rangle, \\ \nabla^2 \text{Scal}, \alpha(\text{Ric}), \\ \langle \text{Ric} \otimes \text{Ric}, \dot{R} \rangle, \langle \text{Ric}, \dot{R} \rangle, \\ \dot{R}, \bar{R}, \ \nabla R\ ^2, \\ \langle \Delta R, R \rangle \end{array} \right\}$ | görbületi invariánsok                  | 46, 47   |
| $R_{(E_i)_{1 \leq i \leq n}}^{(\beta)}$   | Gibbs-féle eloszlás                    | 19   |
| $R_{(H)}^{(\beta)}$   | Gibbs-állapot                          | 89   |
| $R(X, Y)Z$  | görbületi tenzor                       | 38   |
| $\text{Ric}$  | Ricci-féle görbületi tenzor            | 39   |
| $S(D)$  | Neumann-féle entrópia                  | 88   |
| $S(f)$  | az $f$ eloszlás entrópiája             | 14   |
| $S(X, Y)$   | $M \rightarrow \mathbb{R}$ függvény    | 53   |
| $\text{Scal}$   | skalárgörbület                         | 41   |
| $\tilde{\mathcal{P}}_n$   | a $\mathcal{P}_n$ tér bővítése         | 52   |
| $TM$  | érintőnyaláb                           | 34   |
| $T_p M$   | $p$ pontbeli érintőtér                 | 33   |
| $T_D^\perp \mathcal{M}_n^+$   | az állapottér érintőterének altere     | 107  |
| $T_D^C \mathcal{M}_n^+$   | az állapottér érintőterének altere     | 107  |
| $V_\vartheta(\tilde{\vartheta})$  | variancia                              | 13   |
| $V_{D_\vartheta}^{(n), f}$  | $f$ -variancia a $D_\vartheta$ pontban | 106  |
| $V(U)$  | az $U$ tartomány térfogata             | 42   |



# Tárgymutató

- $\alpha$ -
  - kovariáns deriválás, 38
  - relatív entrópia, 23
- ív hossz, 41
- állapot, 86
  - Gibbs, 89
  - kevert, 86
  - legkevertebb, 89
  - tiszta, 86
- állapottér
  - $n$ -dimenziós, 86
  - komplex, 86
  - valós, 86
- általánosított
  - divergencia, 22
  - kontrasztfüggvény, 22
- átmenetvalószínűség, 12
- érintő
  - nyaláb, 34
  - tér, 33
- affinösszefüggő sokaság, 37
- beágyazás, 50
- becslés, 13, 106
  - $f$ -varianciája, 106
  - torzítatlan, 106
- Bhattacharyya-féle relatív entrópia, 23, 77
- Boltzmann
  - $\sim$ -féle transzportegyenlet, 15
  - állandó, 18
  - entrópia, 15
- Bures-metrika, 101
- Cencov–Morozova-féle függvény, 96
- Christoffel-szimbólum, 37
- Cramer–Rao-tétel, 13, 107
- Csiszár-féle divergencia, 24
- deriváció, 33
- derivált leképezés, 50
- differenciálgeometria, 37
  - lapos, 38
  - torziómentes, 37
- differenciálható
  - függvény, 33
  - sokaság, 32
- diszkrét
  - eloszlás, 32
- diszkrét eloszlás, 77
- divergencia, 25
  - $\alpha$ - $\sim$ , 23
  - általánosított, 22
  - Bhattacharyya-féle, 23, 77
  - Csiszár-féle, 24
  - duális, 25
  - háromszög, 23
  - harmonikus, 23
  - Hellinger-féle, 22
  - Jeffreys-féle, 23
  - Kullback–Liebler-féle, 22
  - Lin–Wong-féle, 23
  - példák, 22
  - sorfejtése, 24
- duális divergencia, 25
- Einstein-féle írásmód, 36
- elégéges statisztika, 11
- elfogadható
  - monoton metrika, 164
- eloszlás
  - diszkrét, 8, 32, 77
  - Gibbs-, 16, 19, 28
  - normális, 8, 55, 78
  - egydimenziós, 9, 14
- entrópia, 14, 15, 19
  - maximális  $\sim$  elve, 16, 19

- Neumann-féle, 88  
 események, 83, 85  
   diszjunktak, 85  
   elemi~, 85  
   teljes rendszere, 85  
 exponenciális  
   család, 7, 8  
   leképezés, 44, 45  
   paraméterezés, 7  
 függvény  
   beágyazás, 50  
   Cencov–Morozova-féle, 96  
   csíra, 33  
   derivált leképezése, 50  
   differenciálható, 33  
   duálisa, 95, 113  
   m-monoton, 144  
   majorizációra nézve monoton, 140  
   normált, 113  
   normalizált, 95  
   operátorkonvex, 112  
   operátormonoton, 93  
   sima, 33  
   szimmetrikus, 95, 113  
   transzponáltja, 95, 113  
 felületi merőleges, 53  
 Fisher-féle  
   metrika, 26, 36, 48, 50, 56  
 Fisher-féle információ  
   ~s mátrix, 9, 36, 105  
   additivitása, 12  
   monotonitása, 11  
 fizikai mennyiség, 87  
    $k$ -adik momentuma, 87  
   szórása, 87  
   várható értéke, 87  
 görbület  
   ~i tenzor, 38, 64  
   Ricci-féle, 39, 40, 48  
   skalár~, 41, 54, 64, 74  
   geodetikus, 42, 45, 77–79, 119, 191  
   Gibbs-állapot, 89  
   Gibbs-eloszlás, 16, 19, 28  
   háromszög relatív entrópia, 23  
   hőmérséklet, 18, 89  
   harmonikus  
     relatív entrópia, 23  
   Hellinger-féle relatív entrópia, 22  
   Hilbert–Schmidt-féle skalárszorzás, 58, 94  
   indukált Riemann-metrika, 50  
   információ  
     ~s veszteség, 11, 12  
   Jeffreys-féle relatív entrópia, 23  
   kétszeresen sztochasztikus mátrix, 28  
   kanonikus  
     korreláció, 102  
     paraméterezés, 7, 92  
     skaláris szorzás, 58, 94  
   kevert  
     család, 7, 8  
     paraméterezés, 7  
   kommutátor, 35  
   konnexió, 37  
      $\alpha$ -~, 38  
     Levi–Civita-féle, 38  
     Riemann-~, 37  
   kontrasztfüggvény, 25  
     általánosított, 22  
   koordinátázás  
     lokális, 32  
   koordinátarendszer  
     lokális, 32  
   kovariáns deriválás, 37  
      $\alpha$ , 38  
     vektor szerinti, 37  
     visszahúzottja, 50



- Kullback–Liebler-féle  
relatív entrópia, 22
- kvantummechanikai  
Fisher-féle információ, 105
- Löwner-tétele, 96
- lapos differenciálgeometria, 38
- legkevertebb állapot, 89
- legkisebb monoton metrika, 101
- legnagyobb monoton metrika, 101
- Levi–Civita-féle konnexió, 38
- Lin–Wong-féle relatív entrópia, 23
- logaritmikus derivált  
szimmetrikus, 101
- lokális  
koordinátázás, 32  
koordinátarendszer, 32  
térkép, 32
- majorizáció, 27, 28, 89
- maximális entrópia elve, 16, 19
- metrika  
Bogoljubov-féle, 102  
Bures-féle, 101  
Fisher-féle, 36, 48, 50, 56  
indukált, 50  
Kubo–Mori, 102  
legkisebb monoton, 101  
legnagyobb monoton, 101  
monoton, 95  
elfogadható, 164  
Riemann-~, 36  
visszahúzottja, 50
- minta, 13
- monoton metrikák  
családja, 93
- Neumann-féle entrópia, 88
- normális eloszlás, 8, 55, 56, 78
- normálvektormező, 53
- obszervábilis, 87
- operátor  
balról szorzás  $\sim a$ , 94  
jobbról szorzás  $\sim a$ , 94  
teljesen pozitív, 91
- operátorkalkulus  
Riesz–Dunford-féle, 97
- operátorkonvex függvény, 112
- Pauli-mátrixok, 88
- Petz  
sejtése, 140
- Petz osztályzási tétel, 94
- részsokaság, 50  
1 kodimenziós, 50
- relatív entrópia, 25  
 $\alpha$ -~, 23  
 $g$ -, 113  
Bhattacharyya-féle, 23  
Bures-féle, 117  
Csiszár-féle, 24  
differenciálható, 112  
háromszög, 23  
harmonikus, 23  
Hellinger-féle, 22  
Jeffreys-féle, 23  
Kullback–Liebler-féle, 22  
kvadratikus, 118  
Lin–Wong-féle, 23  
monoton, 112  
példák, 22  
sorfejtése, 24  
szimmetrikus, 112  
szimmetrizáltja, 114  
távolság, 112  
Umegaki-féle, 118
- Ricci-féle görbület, 39, 40, 48
- Riemann  
-konnexió, 37  
geometria, 36  
metrika, 36

- Fisher-féle, 36, 56
- indukált, 50
- sokaság, 36
- Riesz–Dunford-féle kalkulus, 97
- sűrűségi mátrix, 86
- sima
  - függvény, 33
  - sokaság, 32
  - vektormező, 35
- skalárgörbület, 41, 54, 64, 74, 139, 166
- skaláris szorzás
  - Hilbert–Schmidt-féle, 58
  - kanonikus, 58
- sokaság
  - $C^\infty$ -beli, 32
  - 1 kodimenziós rész-, 50
  - affinösszefüggő, 37
  - differenciálható, 32
  - rész-, 50
  - Riemann-~, 36
  - sima, 32
- statisztika
  - elégseges, 11
- statisztika varianciája, 13
- statisztikai
  - becslés, 13
  - minta, 13
- statisztikai modell, 6
  - $f$  általi képe, 10
  - ek szorzata, 12
  - kvantummechanikai, 90
- Stokes-paraméterezés, 87
- sztochasztikus leképezés, 91
  
- T-transzformáció, 28
- távolság, 76
- térfogat, 42, 190
  - Taylor-sora, 192
  - Taylor-sora, 48
- térkép
  - lokális, 32
  - teljesen pozitív operátor, 91
  - torziómentes differenciálgeometria, 37
  - vektormező, 35
    - $\sim$ k kommutátora, 35
    - normál-, 53

## Hivatkozások

- [1] P. M. Alberti, A. Uhlmann. *Stochasticity and partial order*, volume 18 of *Mathematische Monographien [Mathematical Monographs]*. VEB Deutscher Verlag der Wissenschaften, Berlin, 1981.
- [2] S. Amari. *Differential-geometrical methods in statistics*, volume 28 of *Lecture Notes in Statistics*. Springer-Verlag, New York, 1985.
- [3] S. Amari, H. Nagaoka. *Methods of information geometry*, volume 191 of *Translations of Mathematical Monographs*. American Mathematical Society, Providence, RI, 2000.
- [4] Andai A.. On the monotonicity conjecture for the curvature of the Kubo-Mori metric. arXiv:math-ph/0310064.
- [5] Andai A.. Monotone riemannian metrics on density matrices with non-monotone scalar curvature. *J. Math. Phys.*, 44(9):3675–3688, 2003.
- [6] T. Ando. Concavity of certain maps on positive definite matrices and applications to Hadamard products. *Linear Algebra Appl.*, 26:203–241, 1979.
- [7] T. Ando. Majorization, doubly stochastic matrices, and comparison of eigenvalues. *Linear Algebra Appl.*, 118:163–248, 1989.
- [8] W. B. Arveson. Subalgebras of  $C^*$ -algebras. *Acta Math.*, 123:141–224, 1969.
- [9] R. Balian, Y. Alhassid, H. Reinhardt. Dissipation in many-body systems: a geometric approach based on information theory. *Phys. Rep.*, 131(1-2):1–146, 1986.
- [10] O. E. Barndorff-Nielsen. *Parametric statistical models and likelihood*, volume 50 of *Lecture Notes in Statistics*. Springer-Verlag, New York, 1988.
- [11] M. B. Bassat.  $f$ -entropies, probability of error and feature selection. *Inform. Control*, 39:227–242, 1978.
- [12] V. P. Belavkin, P. Staszewski.  $C^*$ -algebraic generalization of relative entropy and entropy. *Ann. Inst. H. Poincaré Sect. A (N.S.)*, 37(1):51–58, 1982.
- [13] J. Bertrand, C. F. Diguët, V. Puiseux. Démonstration d’un théorème de Gauss. *Journal de Mathématiques*, 13:80–90, 1848.
- [14] R. Bhatia. *Matrix analysis*, volume 169 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1997.

- [15] A. Bhattacharyya. On a measure of divergence between two statistical populations defined by their probability distributions. *Bull. Calcutta Math. Soc.*, 35:99–109, 1943.
- [16] Bognár J., Göndöcs F., Kászonyi L., Kováts A., Michaletzky Gy., Móri T., Somogyi Á., Szeidl L., Székely J. G.. *Matematikai statisztika*. Nemzeti Tankönyvkiadó, Budapest, 1995.
- [17] Bolla M., Csiszár I., Gaudi I., Gulyás O., Hajtman B., Kun A., Lengyel T., Michaletzky Gy., Rejtő L., Rudas T., Székely Gy., Telegdi L., Tusnády G.. *Többváltozós statisztikai analízis*. Műszaki Könyvkiadó, Budapest, 1986.
- [18] D. C. Brody, A. Ritz. Geometric phase transitions. cond-mat/9903168, 1999.
- [19] N. N. Čencov. *Statistical decision rules and optimal inference*, volume 53 of *Translations of Mathematical Monographs*. American Mathematical Society, Providence, R.I., 1982.
- [20] Man Duen Choi. Completely positive linear maps on complex matrices. *Linear Algebra and Appl.*, 10:285–290, 1975.
- [21] J. Conway. *A Course in Functional Analysis*. Graduate Texts in Mathematics. Springer Verlag, 1990.
- [22] Csiszár I.. Information-type measures of difference of probability distributions and indirect observations. *Studia Sci. Math. Hungar.*, 2:299–318, 1967.
- [23] Csiszár I.. On topology properties of  $f$ -divergences. *Studia Sci. Math. Hungar.*, 2:329–339, 1967.
- [24] C. Davis. Notions generalizing convexity for functions defined on spaces of matrices. In *Proc. Sympos. Pure Math., Vol. VII*, pages 187–201. Amer. Math. Soc., Providence, R.I., 1963.
- [25] J. Dittmann. On the curvature of monotone metrics and a conjecture concerning the Kubo-Mori metric. *Linear Algebra Appl.*, 315(1-3):83–112, 2000.
- [26] J. Dittmann, A. Uhlmann. Connections and metrics respecting purification of quantum states. *J. Math. Phys.*, 40(7):3246–3267, 1999.
- [27] A. Dvurečenskij. *Gleason's theorem and its applications*, volume 60 of *Mathematics and its Applications (East European Series)*. Kluwer Academic Publishers Group, Dordrecht, 1993.

- [28] B. Efron. Defining the curvature of a statistical problem (with applications to second order efficiency). *Ann. Statist.*, 3(6):1189–1242, 1975.
- [29] E. Fick, G. Sauer mann. *The quantum statistics of dynamic processes*, volume 86 of *Springer Series in Solid-State Sciences*. Springer-Verlag, Berlin, 1990.
- [30] R. A. Fisher. On the mathematical foundations of theoretical statistics. *Phil. Trans. R. Soc., A*, 222:309–368, 1922.
- [31] R. A. Fisher. The statistical utilization of multiple measurements. *Annals of Eugenics*, 8:376–386, 1938.
- [32] R. A. Fisher. *Statistical methods and scientific inference*. 2nd ed., revised. Hafner Publishing Company, New York, 1959.
- [33] B. R. Frieden. Fisher information, disorder, and the equilibrium distributions of physics. *Phys. Rev. A (3)*, 41(8):4265–4276, 1990.
- [34] B. R. Frieden. *Physics from Fisher information*. Cambridge University Press, Cambridge, 1998.
- [35] A. Fujiwara, H. Nagaoka. Quantum Fisher metric and estimation for pure state models. *Phys. Lett. A*, 201(2-3):119–124, 1995.
- [36] A. Fujiwara, H. Nagaoka. An estimation theoretical characterization of coherent states. *J. Math. Phys.*, 40(9):4227–4239, 1999.
- [37] P. Gibilisco, T. Isola. On characterization of dual statistically monotone metrics. math.PR/0303059, 2003.
- [38] P. Gibilisco, T. Isola. Wigner-Yanase information on quantum state space: the geometric approach. math.PR/0304170, 2003.
- [39] P. Gibilisco, T. Isola. Monotone metrics on statistical manifolds of density matrices by geometry of non-commutative  $l^2$ -spaces. *Infin. Dimens. Anal. Quantum Probab. Relat. Top.*, pages 129–140, 2001.
- [40] A. Gray. The volume of a small geodesic ball of a Riemannian manifold. *Michigan Math. J.*, 20:329–344 (1974), 1973.
- [41] A. Gray, L. Vanhecke. Riemannian geometry as determined by the volumes of small geodesic balls. *Acta Math.*, 142(3-4):157–198, 1979.
- [42] F. Hansen, G. K. Pedersen. Jensen’s inequality for operators and Löwner’s theorem. *Math. Ann.*, 258(3):229–241, 1981/82.

- [43] G. H. Hardy, J. E. Littlewood, Pólya Gy.. *Inequalities*. Cambridge Mathematical Library. Cambridge University Press, Cambridge, 1988.
- [44] H. Hasegawa. Dual geometry of the wigner–yanase–dyson information content.
- [45] H. Hasegawa.  $\alpha$ -divergence of the noncommutative information geometry. In *Proceedings of the XXV Symposium on Mathematical Physics (Toruń, 1992)*, volume 33, pages 87–93, 1993.
- [46] H. Hasegawa. Non-commutative extension of the information geometry. In *Quantum communications and measurement (Nottingham, 1994)*, pages 327–337. Plenum, New York, 1995.
- [47] M. Hayashi. Asymptotic estimation theory for a finite-dimensional pure state model. *J. Phys. A*, 31(20):4633–4655, 1998.
- [48] M. Hayashi. Corrigenda: “Asymptotic estimation theory for a finite-dimensional pure state model”. *J. Phys. A*, 31(41):8405, 1998.
- [49] S. Helgason. *Differential geometry, Lie groups, and symmetric spaces*, volume 34 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2001.
- [50] E. Hellinger. Neue Begründung der Theorie quadratischer Formen von unendlich vielen Veränderlichen. *J. für reine and Angew. Math.*, 36:210–271, 1909.
- [51] F. Hiai, Petz D.. *The semicircle law, free random variables and entropy*, volume 77 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 2000.
- [52] H. Hotteling. Tubes and spheres in  $n$ -spaces, and a class of statistical problems. *Amer. J. Math.*, 61:440–460, 1939.
- [53] R. S. Ingarden, H. Janyszek, A. Kossakowski, T. Kawaguchi. Information geometry of quantum statistical systems. *Tensor (N.S.)*, 37(1):105–111, 1982.
- [54] H. Jeffreys. An invariant form for the prior probability in estimation problems. *Proc. Roy. Soc. London. Ser. A.*, 186:453–461, 1946.
- [55] A. Jenčová. Quantum information geometry and standard purification. *J. Math. Phys.*, 43(5):2187–2201, 2002.
- [56] S. Kobayashi, K. Nomizu. *Foundations of differential geometry. Vol. I,II*. Wiley Classics Library. John Wiley & Sons Inc., New York, 1996.

- [57] K. Kraus. *States, effects, and operations*, volume 190 of *Lecture Notes in Physics*. Springer-Verlag, Berlin, 1983.
- [58] F. Kubo, T. Ando. Means of positive linear operators. *Math. Ann.*, 246(3):205–224, 1979/80.
- [59] R. A. Leibler, S. Kullback. On information and sufficiency. *Ann. Math. Statistics*, 22:79–86, 1951.
- [60] S. Kullback. *Information theory and statistics*. Dover Publications Inc., Mineola, NY, 1997.
- [61] L. D. Landau, E. M. Lifsic. *Elméleti fizika III*. Tankönyvkiadó, Budapest, 1981.
- [62] L. D. Landau, E. M. Lifsic. *Elméleti fizika V*. Tankönyvkiadó, Budapest, 1981.
- [63] S. Lang. *Differential manifolds*. Springer–Verlag, New York, second edition, 1985.
- [64] A. Lesniewski, M. B. Ruskai. Monotone Riemannian metrics and relative entropy on noncommutative probability spaces. *J. Math. Phys.*, 40(11):5702–5724, 1999.
- [65] J. Lin, S. K. M. Wong. A new directed divergence measure and its characterization. *Int. J. General Systems*, 17:73–81, 1990.
- [66] G. Lindblad. Entropy, information and quantum measurements. *Comm. Math. Phys.*, 33:305–322, 1973.
- [67] G. Lindblad. Expectations and entropy inequalities for finite quantum systems. *Comm. Math. Phys.*, 39:111–119, 1974.
- [68] G. Lindblad. Completely positive maps and entropy inequalities. *Comm. Math. Phys.*, 40:147–151, 1975.
- [69] K. Löwner. Über monotone Matrixfunctionen. *Math. Z.*, 38:177–216, 1934.
- [70] A. W. Marshall, I. Olkin. *Inequalities: theory of majorization and its applications*, volume 143 of *Mathematics in Science and Engineering*. Academic Press Inc. [Harcourt Brace Jovanovich Publishers], New York, 1979.
- [71] Marx Gy.. *Kvantummechanika*. Második, átdolgozott és bővített kiadás. Műszaki Könyvkiadó, Budapest, 1964.
- [72] M. Mei. The theory of genetic distance and evaluation of human races. *Japan J. Human Genetics*, 23:341–369, 1978.

- [73] P. W. Michor, Petz D., Andai A.. On the curvature of a certain Riemannian space of matrices. *Infin. Dimens. Anal. Quantum Probab. Relat. Top.*, 3(2):199–212, 2000.
- [74] E. A. Morozova, N. N. Chentsov. Markov invariant geometry on state manifolds. In *Current problems in mathematics. Newest results, Vol. 36 (Russian)*, Itogi Nauki i Tekhniki, pages 69–102, 187. Akad. Nauk SSSR Vsesoyuz. Inst. Nauchn. i Tekhn. Inform., Moscow, 1989.
- [75] M. K. Murray, J. W. Rice. *Differential geometry and statistics*, volume 48 of *Monographs on Statistics and Applied Probability*. Chapman & Hall, London, 1993.
- [76] J. von Neumann. Thermodynamik quantenmechanischer Gesamtheiten. *Gött. Nachr.*, pages 273–291, 1927.
- [77] M. Ohya, Petz D.. *Quantum entropy and its use*. Texts and Monographs in Physics. Springer-Verlag, Berlin, 1993.
- [78] M. Ohya, Petz D.. Notes on quantum entropy. *Studia Sci. Math. Hungar.*, 31(4):423–430, 1996.
- [79] Petz D.. Quasi-entropies for finite quantum systems. *Rep. Math. Phys.*, 23(1):57–65, 1986.
- [80] Petz D.. On certain properties of the relative entropy of states of operator algebras. *Math. Z.*, 206(3):351–361, 1991.
- [81] Petz D.. Geometry of canonical correlation on the state space of a quantum system. *J. Math. Phys.*, 35(2):780–795, 1994.
- [82] Petz D.. Monotone metrics on matrix spaces. *Linear Algebra Appl.*, 244:81–96, 1996.
- [83] Petz D.. Information-geometry of quantum states. In *Quantum probability communications, QP-PQ, X*, pages 135–157. World Sci. Publishing, River Edge, NJ, 1998.
- [84] Petz D.. Covariance and Fisher information in quantum mechanics. *J. Phys. A*, 35(4):929–939, 2002.
- [85] Petz D.. Monotonicity of quantum relative entropy revisited. *Rev. Math. Phys.*, 15(1):79–91, 2003.



- [86] Petz D., H. Hasegawa. On the Riemannian metric of  $\alpha$ -entropies of density matrices. *Lett. Math. Phys.*, 38(2):221–225, 1996.
- [87] Petz D., M. B. Ruskai. Contraction of generalized relative entropy under stochastic mappings on matrices. *Infin. Dimens. Anal. Quantum Probab. Relat. Top.*, 1(1):83–89, 1998.
- [88] Petz D., Sudár Cs.. Geometries of quantum states. *J. Math. Phys.*, 37(6):2662–2673, 1996.
- [89] Petz D., Tóth G.. The Bogoliubov inner product in quantum statistics. *Lett. Math. Phys.*, 27(3):205–216, 1993.
- [90] E. C. Pielou. *Ecological diversity*. Wiley, New York, 1975.
- [91] C. R. Rao. Information and accuracy attainable in the estimation of statistical parameters. *Bulletin of the Calcutta Mathematical Society*, 37:81–91, 1945.
- [92] C. R. Rao. Diversity and dissimilarity coefficients: a unified approach. *Theoretic Population Biology*, 21:24–43, 1982.
- [93] G. Ruppeiner. Riemannian geometry in thermodynamic fluctuation theory. *Rev. Modern Phys.*, 67(3):605–659, 1995.
- [94] M. B. Ruskai. Beyond strong subadditivity? Improved bounds on the contraction of generalized relative entropy. *Rev. Math. Phys.*, 6(5A):1147–1161, 1994.
- [95] K. Sailer. *Nemegyensúlyi statisztikus fizika*. Kossuth Lajos Tudományegyetem, Debrecen, 1994.
- [96] M. Spivak. *A comprehensive introduction to differential geometry. Vol. I, II, III, IV, V*. Publish or Perish Inc., Wilmington, Del., second edition, 1979.
- [97] W. F. Stinespring. Positive functions on  $C^*$ -algebras. *Proc. Amer. Math. Soc.*, 6:211–216, 1955.
- [98] R. F. Streater. Statistical dynamics and information geometry. In *Geometry and nature (Madeira, 1995)*, volume 203 of *Contemp. Math.*, pages 117–131. Amer. Math. Soc., Providence, RI, 1997.
- [99] Szenthe J.. *Bevezetés a sima sokaságok elméletébe*. ELTE Eötvös Kiadó, Budapest, 2002.
- [100] Szolcsányi E.. *Differenciálgeometria I*. Tankönyvkiadó, Budapest, 1991.

- [101] T. Tanaka. Information geometry of mean-field approximation. In *Advanced mean field methods (Birmingham, 1999)*, Neural Inf. Process. Ser., pages 259–273. MIT Press, Cambridge, MA, 2001.
- [102] I. J. Taneja. Generalised information measures and their applications. preprint. <http://www.mtm.ufsc.br/~taneja/bhtml/bhtml.html>.
- [103] F. Topsøe. Some inequalities for information divergence and related measures of discrimination. *Res. Rep. Coll., RGMIA*, 2(1):85–98, 1999.
- [104] A. Uhlmann. Relative entropy and the Wigner-Yanase-Dyson-Lieb concavity in an interpolation theory. *Comm. Math. Phys.*, 54(1):21–32, 1977.
- [105] A. Uhlmann. Geometric phases and related structures. In *Proceedings of the XXVII Symposium on Mathematical Physics (Toruń, 1994)*, volume 36, pages 461–481, 1995.
- [106] A. Uhlmann. Spheres and hemispheres as quantum state spaces. *J. Geom. Phys.*, 18(1):76–92, 1996.
- [107] H. Umegaki. Conditional expectation in an operator algebra. IV. Entropy and information. *Kōdai Math. Sem. Rep.*, 14:59–85, 1962.
- [108] V. S. Varadarajan. *Geometry of quantum theory. Vol. I*. D. Van Nostrand Co., Inc., Princeton, N.J.-Toronto, Ont.-London, 1968.
- [109] V. S. Varadarajan. *Geometry of quantum theory. Vol. II*. Van Nostrand Reinhold Co., New York, 1970.
- [110] H. Vermeil. Notiz über das mittlere Krümmungsmass einer  $n$ -fach ausgedehnten Riemann'schen Mannigfaltigkeit. *Akad. Wiss. Göttingen Nachr.*, pages 334–344, 1917.
- [111] A. Wehrl. How chaotic is a state of a quantum system? *Rep. Math. Phys.*, 6:15–28, 1974.
- [112] E. P. Wigner, Mutsuo M. Yanase. Information contents of distributions. *Proc. Nat. Acad. Sci. U.S.A.*, 49:910–918, 1963.
- [113] W. K. Wothers. Statistical distance and Hilbert space. *Physical Review D*, 23:357–362, 1981.